DOCUMENT RESUME

ED 304 504 UD 026 685

AUTHOR Yagi, Kan; Kushman, James W.
TITLE Longitudinal Study of Students In and Out of a Chapter 1 Program over a Three-Year Period in the Portland Public Schools.
INSTITUTION Portland Public Schools, OR. Research and Evaluation Dept.
PUB DATE Nov 88
NOTE 62p.
PUB TYPE Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS Academic Achievement; *Achievement Gains; Achievement Tests; *Basic Skills; *Classroom Environment; *Compensatory Education; Educational Environment; Elementary Education; *Elementary School Students; Longitudinal Studies; Program Evaluation; Student Behavior; Supplementary Education
IDENTIFIERS *Education Consolidation Improvement Act Chapter 1; Portland Basic Skills Achievement Levels Testing; *Portland School District OR

ABSTRACT

Students in the Portland (Oregon) Public Schools' Education Consolidation Improvement Act Chapter 1 program made greater achievement gains while in the program than when they were not, and those gains are probably attributable to the extent that Chapter 1 students engage in positive behaviors rather than to the curricular quality of the programs. Chapter 1 students had been observed to consistently engage in socially acceptable and educationally worthwhile classroom behavior. It was hypothesized that this positive behavior would have a significant effect on student achievement gains. Patterns of the 3-year achievement score gains on the Portland Achievement Level Tests for students who were in and out of Chapter 1 programs were displayed on graphs. Two different 3-year periods, 1983-86 and 1984-87, were included. Results were expressed in terms of deviations above or below the gains expected for comparable students. The results were examined for the extent of agreement between the predicted and the obtained patterns displayed in the graphs. The results showed that the gains were usually noticeably larger than the expected when the students were in Chapter 1, and usually noticeably smaller than the expected when they were not. Statistical data are included on 28 graphs and four tables. The appendix contains a comparison of the predictions of the participants' performance with their actual achievement in the reading and the mathematics sequences. A list of 16 references is included. (FMW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

3-20-89

ED304504

1987-1988 EVALUATION REPORT

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Kan Yagi
Portland Public Schools

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

LONGITUDINAL STUDY OF STUDENTS IN AND OUT OF A CHAPTER 1 PROGRAM OVER A THREE-YEAR PERIOD

IN THE PORTLAND PUBLIC SCHOOLS

Research and Evaluation Dept.
Portland Public Schools
Portland, Oregon
Walter E. Hathaway, Director

Kan Yagi
James W. Kushman

November, 1988

UD 026 685

BEST COPY AVAILABLE

LONGITUDINAL STUDY OF STUDENTS IN AND OUT OF A CHAPTER 1 PROGRAM OVER A THREE YEAR PERIOD

KAN YAGI
JAMES W. KUSHMAN

RESEARCH AND EVALUATION DEPARTMENT
PORTLAND PUBLIC SCHOOLS
PORTLAND, OREGON

NOVEMBER 1988

ADMINISTRATIVE SUMMARY

LONGITUDINAL STUDY OF STUDENTS IN AND OUT OF A CHAPTER 1 PROGRAM OVER A THREE YEAR PERIOD 1988-89

Basic skills achievement scores of Chapter 1 students are a result of their experiences not only in Chapter 1 classes but in regular classrooms, in other school rooms, and outside of school. The purpose of the study was to test hypotheses about the effect of student behaviors on gains in academic achievement. Specifically, variations in the extent students engaged in socially acceptable and educationally worthwhile behavior were observed among schools, within classrooms, and, of course, among students. In Chapter 1, however, this behavior was observed to be quite consistent among participating students and classrooms. This behavior and the fact that it appeared to be controllable was expected to be particularly significant especially for students who were likely to be enrolled in a Chapter 1 program at one time or another. It was expected to have a direct effect on basic skills achievement scores as measured by the Portland Achievement Levels Tests. Obviously, ability levels affect student achievement scores; however, the magnitude of the effect was expected to be considerably less than the effect of their behavior on reaching an acceptable level of mastery of a subject. The postulation about socially acceptable and educationally worthwhile behavior, hereafter alternately referred to as positive or negative behavior, takes into account all student behaviors that occur anywhere on school premises. The basic hypothesis was that positive behavior (behavior that was both socially acceptable and educationally worthwhile) would have a noticeably large effect on student achievement gains. In general, the gains would be noticeably larger than average when they were in Chapter 1, and noticeably smaller than average when they were not.

Due to the nature of the behaviors being studied, and the paradigms of student participation/non-participation in Chapter 1 necessary to test the basic hypothesis, the only research design that appeared to be satisfactory was one that incorporated hypotheses about positive/negative behaviors and predictions of student achievement scores which the behaviors were expected to effect. Variations in Fall-Spring achievement gain scores of students who were enrolled in Chapter 1 at least one year of their three-year period of enrollment in a Chapter 1 school were predicted. Seven different groups of enrollment sequences were possible: one group enrolled in Chapter 1 all three years, three enrolled for two of three years, and three enrolled for one of three years. The study involved students in grades three through eight, so each student would be in only one of four grade sequence groups, i.e., grades 3,4,5; 4,5,6; 5,6,7, or 6,7,8. Variations in the patterns of achievement score gains were predicated largely on (1) Chapter 1 participation in programs which were assumed to be much the same

across schools in terms of the key variable, and (2) the strength of the positive/negative behavior tendencies of students in each group.

The study was essentially a replication of one that was conducted four years earlier with refinements. From electronically stored data, patterns of three-year achievement score gains for students who were in and out of Chapter 1 programs in the seven different sequences were displayed in graphs. Like the earlier study, the present study included two different three-year periods, 1983-1986 and 1984-1987, which constituted a replication. Also, Chapter 1 had reading and math programs, so separate sets of graphs were used to display the results by three-year periods and subjects. Another replication was the four grade sequence groups. Results were expressed in terms of deviations above or below the gains expected for comparable students. The results were examined by the researchers, and simple judgments of the extent of agreement between the predicted and obtained patterns displayed in graphs were made. Fall scores obtained by each sequence group at the beginning of the three-year period were also expected to vary in a predictable way.

The results showed that the gains were usually noticeably larger than the expected when the students were in Chapter 1, and usually noticeably smaller than the expected when they were not. Variations in amounts of gains across groups were also predicted. The obtained gain score patterns compared well to those predicted for each sequence group in terms of direction and magnitude. Beginning Fall scores were essentially in the same rank order as they were predicted.

It is contended that the accurate prediction of achievement score patterns demonstrates a causal relationship of socially acceptable and educationally worthwhile behavior to achievement gain scores. It is further contended that given evidence of the significance and effect of student behavior, school personnel, especially regular classroom teachers, would be better able to design ways to cope with this problem than they would to implement most of the other effective practices currently noted in research literature.

TABLE OF CONTENTS

		Page
1.	Introduction	1
	a. Background	1
	b. Purpose and Rationale	4
2.	Method	13
	a. Study Design	13
	b. Subjects	15
	c. Independent and Dependent Variables	16
	d. Predictions/Hypotheses	18
3.	Results	23
4.	Discussion	25
5.	Figures 1-7	30
6.	Tables 1-4	37
7.	References	41
8.	Appendix	43

LONGITUDINAL STUDY OF STUDENTS IN AND OUT OF A CHAPTER 1 PROGRAM OVER A THREE YEAR PERIOD

The purpose of this study was to address a number of the most vexing questions common to Chapter 1 programs in the Portland Public Schools (PPS). Are Chapter 1 programs working? What works and what doesn't? How can the program be improved? Why are some students in the program year after year? Do the students sustain their gains? During deliberations over an appropriate paradigm for sustained gains studies, it became apparent that some intriguing information might be obtained simply by asking the right questions, making the right assumptions, and interrogating electronically stored data in the right way. There was no need to design an elaborate research study, at least not at this time. If the research strategy as envisioned were successful in producing the expected results, it would represent a break from traditional research studies done in school settings where inferences are based largely on results from correlational studies. The present study is the first step to test a theory of student behavior--a step that is necessary to determine if future research on the theory is worthwhile. Subsequent studies of the theoretical variable can then be designed and undertaken with greater confidence in achieving the desired results.

Background. Since its inception over 20 years ago, Title I, now Chapter 1 has been the subject of hundreds of articles. An ERIC search of over 500 ESEA or ECIA articles identified as either evaluation or research reports and conference papers received through 1987 was narrowed to 338. Abstracts, some titles, and 30 complete articles were reviewed in the search for experimental or quasi-experimental studies designed to examine relationships between treatments and outcomes or between independent and dependent variables. Early articles through the seventies were

largely comments and descriptions of aspects of ESEA programs and a much smaller number was largely reports of achievements of participants in the programs defined simply as Title I or by product name.

In the eighties the articles that began to appear were district or state evaluation reports along with a continued number of comments and descriptions of aspects of the law, particularly ECIA, Chapter 1 due to the change in the law. There began to appear reports of student achievements as they participated in various Chapter 1 programs, some of which were identified by instructional system or product. In spite of the large number of articles that were identified as research reports or conference papers which are often based on research, there were few like those of Whitmer (1986) and Yap (1987) that met the standard of having adequate control or same-group achievement data along with experimentally defined and well controlled study variables. The virtual absence of well designed research studies involving Chapter 1 students to demonstrate causal relationships between inputs and outputs is surprising given the large number of articles that purport to have answers to the key questions identified above. Even though Chapter 1 programs are being encouraged to implement the findings of effective schools research, there is no clear evidence that the practices will be effective with Chapter 1 students. The research has only demonstrated relationships, generally low to moderate, and causal relationships are inferred. Nowhere has it been demonstrated which practices or combination of practices will consistently be effective with Chapter 1 students.

Another characteristic of research to date is the focus on treatment, e.g., computerized instruction, rewards and punishment, special reading programs, and not on modifiable student behaviors that affect student performances in general. Some studies on time and time on task like those of Good and Beckerman (1978), Kielsing

(1975), and Wyne and Stuck (1979) come close, but perhaps in the effort to apply measurement to the variable, i.e., clock time, the research suffers in not taking into account total student learning time.

Rigorous scientific investigation generally requires that measures of at least the key study variables be as accurate as possible. In education it is often impractical, if not impossible, to measure some things, but researchers continue to try even at the expense of using far less than satisfactory measuring instruments and research designs, e.g., rating scales, correlational designs, that do not enable them to prove a point. Perhaps we have forgotten that prediction of outcomes is a goal of science and should be a major goal in educational research. The research strategy employed in the study to be reported here uses assumptions, estimates, guesses, and, at best, relative measures of the key study variable, and predictions of outcomes which are measured.

The initial test of the research strategy to demonstrate a causal relationship between an identified experimental variable and student achievement was the target of an earlier study by Yagi (1984). The study spanned the school years from 1979-80 to 1982-83. The present study is a refinement and replication of this earlier study and a further test of the research strategy that involved the prediction of patterns of achievement gains of groups of students who were in and out of Chapter 1 programs over a three year period. This later study spanned the school years from 1983-84 to 1986-87. In these studies, predictions were tied to assumptions about student behaviors which can be controlled largely in schools. Purportedly, if achievement gains of students over the three years could be predicted accurately, the study would not only provide information on the questions identified above, but just as importantly, the study would shed some light on what makes a Chapter 1 student a Chapter 1 student!

In the earlier study, the study groups consisted of seven possible combinations or patterns of enrollment/non-enrollment in Chapter 1 over a three year period; one group of students enrolled in Chapter 1 all three years, three groups of students enrolled for two of three years, and three of students enrolled in Chapter 1 for only one of the three years. Because the study spanned four years, the two different sets of students that were included constituted two replications. The present study reported here also spanned four years constituting two replications. Claus, Girrbach, and Quimper (1982) and DeVito and Rubenstein (1977) also examined the performances of combinations of groups of students who were in and out of Chapter 1; however, they simply followed students for one or two years after their initial treatment year, and they did not speculate a priori on the outcomes.

The enrollment of several thousand students in Chapter 1 each year, or data accumulated over several years, is necessary to meet sample size requirements for the seven study groups in the research design, the reason for the replications. PPS is a large city school district that enrolls over 50,000 students in 80 elementary and middle schools, 10 high schools, and a number of special schools and programs. Each year, about 50 elementary schools and 6 high schools have Chapter 1 programs that enroll over 7,000 students in supplementary reading and/or math programs. The population of Chapter 1 students was expected to be large enough to make the research strategy feasible; but, to be sure, replications with two three-year periods were used.

Purpose and Rationale. To ask the right questions, make the right assumptions, and make the right predictions required insight into how PPS schools and Chapter 1 programs operated and what students in these schools were like. First, essentially all PPS Chapter 1 reading and math programs were pull-out; a critical consideration

because with this arrangement Chapter 1 students would be exposed to an additional period of reading or math instruction each day. Obviously, student achievement scores reflect the students' total school experiences. Here it was assumed that the amount students gained through experiences in Chapter 1 would be added to the amount they gained in the regular classroom. Thus, knowledge of student experiences in regular classrooms as well as Chapter 1 was essential. Knowledge of what most Chapter 1 programs were like was acquired during visits to the various programs for ongoing evaluation and other research study purposes. Information about student experiences in regular classrooms that was obtained during these observations was deemed sufficient for the purposes of this study. The first step was to verify that PPS Chapter 1 programs contributed in some predictable way.

Like schools throughout the nation, PPS schools designed their own Chapter 1 programs, each of which ended up being quite unique despite the fact that a number of schools used the same, commercially designed and marketed systems. PPS studies involving schools using these commercial systems showed no consistent differences among them and no differences when compared to programs designed by individual schools (Adams, Ingebo, Leitner, & Miller, 1983; Adams, Ingebo, Leitner, Miller, Moilanen, and Peters, 1983; Ingebo, Yagi, Leitner, Miller, & Miller, 1982; Kushmuk, 1985a, 1985b; Leitner & Ingebo, 1984a, 1984b). These uniquenesses, along with a number of other conditions which could not be accounted for, posed difficulties in attempting to obtain findings that would allow the researchers to generalize outcomes obtained from a particular system to other schools, grades, subjects, etc. Thus, it seemed necessary to make a general assumption about the Chapter 1 programs which, if correct, would support findings that would be hypothesized in the present study. The assumption was that the achievements of Chapter 1 students in particular kinds of Chapter 1 instructional programs would essentially be equal, and accounting

for any differences for study groups attributable to a particular kind of program was not necessary.

Because is was not feasible to know what instruction and student experiences in all other subjects in Chapter 1 schools was like, it was necessary to rely upon impressions and experiences acquired incidentally from routine visits and sample observations in schools as bases for assumptions. It was assumed that the magnitude of the achievement gains of Chapter 1 students were not only affected by their experiences in the supplementary Chapter 1 reading or math programs, but also by their experiences in all subjects in their regular classrooms including reading and math. With regard to instruction across all subjects, the assumption was like that for Chapter 1 programs, i.e., that reading and math achievements of Chapter 1 students attributable to regular classes would essentially be equal for all study groups, and accounting for any study group differences attributable to particular school programs, again, was not necessary.

It would have been simple to predict that the students would make larger gains when they were in Chapter 1 than when they were not. Claus, Girrbach, and Quimper (1982) and DeVito and Rubenstein (1977) found this for groups of students whose Chapter 1 experience was "interrupted" by a year without Chapter 1. But, would the gains be equal for participating or non-participating years regardless of what year during the three year period the students were in, or out, of Chapter 1? How accurately could the magnitude of the gains be predicted? At this point a return to the critical questions is important.

With regard to the question of what works and what doesn't, i.e., what is effective and what is not, the interest was, and continues to be, on those practices that would be most efficiently and effectively implemented in the District. For example, suggestions

to change instructional materials or classroom arrangements are more easily accepted than suggestions to change instructional styles or parent practices. Thus, the study would not only show if Chapter 1 programs were working, but, hopefully, would reveal something about what made them work.

The question of whether students sustained their gains deserves special consideration not only because federal regulations required that sustained gain studies be conducted, but because of obvious educational significance. Presumably, by using the suggested paradigms of analyzing test data taken at three different times at intervals up to three years, one could determine to what extent gains were sustained. Nowhere was it suggested that consideration must be given to Chapter 1 experience following the initial treatment. Thus, what in particular **causes** the gains to be sustained, or not sustained, is left to speculation. Would not the gains of students be different if during the next year they were again in Chapter 1 or if they were not? How much different might the gains be and why? How much gain should students be expected to sustain? A more important questions is what are the conditions that would maximize and sustain the gains? When conditions are optimal and practical for a school to maintain, only then should an answer to the question of how much to be expected be sought.

As mentioned earlier, the magnitude of the achievement gains of Chapter 1 students were not only affected by their experiences in the supplementary Chapter 1 reading or math programs, but by all their experiences in school including reading and math in their regular classrooms. Based on incidental observations of students in various school circumstances and student/teacher interactions in classrooms, general assumptions were made about the behavior of students in Chapter 1, in their regular classrooms and elsewhere during school hours. First, consider experiences in Chapter 1.

Instruction in PPS Chapter 1 classrooms was individualized or small group seldom with more than ten students in the group or classroom at a time. Teachers and aides usually worked with students individually or in groups up to three or four. Students were nearly always engaged in learning during the time this direct instruction was being provided and their behavior was generally monitored when they worked on their own. Although there were noticeable differences in the kinds and qualities of Chapter 1 supplementary instructional programs, there was little to suggest which curricular or instructional arrangement would consistently result in better performances. In other words, variations in the amount of gain among students attributable to instructional materials or methods used in a program could not be predicted. But, because the amount of time students spent in supplementary reading or math instruction was about equal to the time they also spent in regular classrooms in these subjects, the gains for students with Chapter 1 could be noticeably larger than the average gain for students in their grade. That is, their gains in Chapter 1 would be added to their gains through all other school experiences. Some idea of how much these latter gains might be was needed.

The behavior of every student is not socially acceptable or educationally worthwhile at all times. That is, they successfully avoid learning or engage in socially unacceptable or non-educationally worthwhile behavior in amounts that vary among students. The term "socially acceptable and educationally worthwhile behavior" is used here to distinguish it from terms such as "on task" or "engaged time" which are more familiar, but have different meanings among researchers. It is argued that in order for learning which has educational significance to occur, the behavior of the individual must be both socially acceptable and educationally worthwhile. The terms are used rather broadly and, of necessity, depend on judgment and inferences for definition because the behavior may be covert as well as overt and the motives

may be unconscious as well as conscious. Perhaps some examples will help to clarify their meaning.

During seatwork, a student needlessly goes to the pencil sharpener several times, each time stopping to talk to a friend which obviously interrupts the friend's work. This behavior is neither socially acceptable nor educationally worthwhile. A student sitting quietly and looking attentively at the teacher while instructions are being given is said to be engaged in socially acceptable and educationally worthwhile behavior. The same behavior might be socially acceptable but not educationally worthwhile if the student's eyes wander around the room or outside, and the student pays no attention to what the teacher is saying. Conversely, the behavior of a student that takes a book away from another in order to read it may be educationally worthwhile but not socially acceptable. In the last two examples, the negative behaviors "cancel out" the positive, so the experience has far less educational significance than if all behaviors were positive.

The behaviors extend beyond the classroom into hallways, cafeterias, restrooms, and on playgrounds and elsewhere. For example, the behavior of a student who barges in at the head of a cafeteria line is neither socially acceptable nor educationally worthwhile. Or, the behavior of a student who vigorously engages in a game at the expense of another may be educationally worthwhile but not socially acceptable. Students who consistently engage in mischievous behavior in the classroom are likely to engage in such behavior elsewhere. Conversely, students who consistently engage in socially acceptable and educationally worthwhile behavior in the classroom are likely to engage in such behavior elsewhere. For convenience the general terms "positive" or "negative" behaviors will sometimes be used when both socially acceptable and educationally worthwhile behaviors, or the opposite, are being referenced.

The extent of observed negative behavior varied among schools and, obviously, among students from a little to a lot. Tracking and measuring this behavior of students for study purposes as the study might have been designed would have been costly and impractical for two reasons which will be clear later: (1) it would have been necessary to observe each student in all circumstances over the duration of the study, and (2) the identity of students in Chapter 1 in later years but not earlier ones could not be known at the outset. Measurement of the key behaviors also seemed unnecessary at this preliminary stage, so the decision was to formulate hypotheses about them, and then use predictions of patterns of student achievement gains to test the hypotheses. If the patterns could be predicted, the contention of hypothesized causal relationships between positive/negative behaviors and achievement gain are supported and future research can be better directed. The two general hypotheses were (1) that the more students engaged in negative behaviors, the more likely they would be enrolled in Chapter 1 and enrolled for a longer time, and (2) when students were enrolled in Chapter 1 their achievement gains would be noticeably larger than when they were not, and, when they were not enrolled in Chapter 1 their gains would not only be smaller but generally much smaller than the average of their peers. In other words, the propensity for PPS Chapter 1 students to engage in negative behavior was to a considerable extent under their control, i.e., it depended on their perceptions of what they could get away with, and that the behavior persisted so long and to the extent it was tolerated. This behavior was expected to affect achievement gains accordingly. Implications of negative behaviors are far reaching. Likewise, the implications for Chapter 1 which has received, and is still receiving, heavy criticism for disappointing performances of low achieving students are far reaching.

Hypotheses about the relationships between student behavior and their performances as measured by reading and math achievement tests are based on the following assumptions:

1. Teachers vary in the amount of negative behavior that is tolerated in classrooms and elsewhere. General tolerated levels are more characteristic of the school than of individual teachers.

2. Some students have a propensity for more negative behavior than others, and unless extinguished, the behaviors will persist across the grades.

3. Students whose behaviors are more negative are likely to obtain lower basic skills achievement test scores and are more likely to be enrolled in Chapter 1 at some time. When students are enrolled in Chapter 1, the proportion of positive behavior, especially educationally worthwhile behavior, increases dramatically.

4. Student differences in basic academic ability, e.g., intelligence, are not ignored, but these differences are assumed to be minor compared to influences of their behavior in mastering the curriculum of supplementary and regular instruction.

5. PPS Chapter 1 instruction is largely pull-out and the key feature is individual attention which tends to keep students engaged in positive behaviors during this time. Negative behaviors are difficult to extinguish and will remain relatively unchanged for most students for the duration of the study.

6. Variations in the amount of time students are engaged in positive behavior is a unique characteristic of the individual schools. These variations among Chapter 1 schools are assumed to be normally distributed within the variations among all schools in the District.

7. The effects of student behavior applies equally to achievement scores in reading and math, and any differences in score patterns for these subjects will be constant.

If the assumptions identified above are accurate, then patterns of achievement gains for groups of students who were in and out of Chapter 1 programs at different times over three successive years ought to be different and predictable. The study was designed to test the hypothesis that

> variations in Chapter 1 student achievement gains as observed over a three year period are predictable, and to a considerable degree, these variations are attributable to the time the students engage in socially acceptable and educationally worthwhile behavior. Chapter 1 contributes substantially to the proportion of time these students engage in these behaviors.

The present study as well as the earlier one was predicated on two fundamental premises: (1) that very little academic achievement as measured by the reading and math basic skills tests can obtain unless students engage in socially acceptable and educationally worthwhile behaviors, and (2) that engagement in these positive behaviors can be efficiently and effectively controlled within schools.

METHOD

Study Design--The general design for the study was to track same-student groups over a three year period measuring Fall-to-Spring achievement gains as a function of participation/non-participation in Chapter 1 supplementary instruction. In the PPS district wide testing program, students in grades three through eight are given the Portland Achievement Levels Tests (PALT) in reading, math, and language usage in both Fall and Spring. Student achievement is measured by RIT (Rasch Unit) scores which are derived from a single, equal interval scale underlying the PPS curriculums which span the six grades in these three subjects (See Parent-Teacher Guide to Portland Achievement Levels Tests, Fall 1987 Edition). These scores are curriculum based in a way that makes them different from commonly used population based scores, e.g., scale scores. RIT scores currently range from 140 to 280 for the three subjects. PPS Chapter 1 programs are for reading and math, so the present study uses only the achievement gains in these subjects to follow student progress over the three year interval.

The longitudinal study design has several advantages over the typical "snapshot" Chapter 1 evaluation which looks at student achievement for a single year. First, the Chapter 1 student groups serve as their own "controls" over time. In the single-year evaluation, finding a suitable comparison group from the same school or another school is difficult. Peers from the same school are likely to be higher ability students while low ability students from non-Chapter 1 schools have different family background characteristics, and besides, may be receiving supplementary instruction under a different program label. On the other hand, student groups from the same Chapter 1 schools who differ only by grade level or Chapter 1 participation/non-participation from year to year are likely to have the same, or nearly the same, ability and family background. Over three years, seven participation

sequences are possible. For example, a "no-yes-no" or N-Y-N sequence (Participating in Chapter 1 the second year but not the first and third) allows a comparison of one treatment year with non-treatment years before and after on the same group of students.

A second advantage of the longitudinal approach is that with the seven participation sequences, one can hypothesize a priori what achievement gain patterns should be like if the assumptions about Chapter 1 supplementary instruction, regular classroom instruction, and student learning behavior are correct. Furthermore, there are four possible grade sequences within each participation sequence (grades 3-4-5, 4-5-6, 5-6-7, and 6-7-8) and two subjects, reading and math. If results fit the predicted patterns consistently over different grade sequences and for both subjects, this is strong evidence that the key condition of socially acceptable and educationally worthwhile behavior characteristic of Chapter 1 is having its predicted effects.

A third advantage is that certain participation sequences (Y-Y-N; N-Y-N) allow one to assess sustained gains. For example, in a N-Y-N sequence, one can look at achievement gains for one baseline year, one program year, and one follow-up year. Student performances in the first follow-up year is an indication of sustained effects. Gains should not drop dramatically to the baseline level if Chapter 1 and regular classroom instruction together are having long-term or sustained effects.

A final advantage is that students in each participating sequence are expected to have different scores at the beginning of the first year which indicates something about why they are enrolled in Chapter 1 the particular year. For example, students enrolled all three years, Y-Y-Y, are expected to have more difficulties, e.g., lower ability, less disciplined, than students in Chapter 1 only during the third year. Therefore, compared to other groups, the

Y-Y-Y groups are expected to have lower initial scores and make smaller gains during the years the groups are in Chapter 1. To test this hypothesis requires that students be enrolled in a school with a Chapter 1 program for all three years. There is no way to know when students are first enrolled how long they will be in Chapter 1; however, if a probable cause is known, something might be done to change the situation.

A distinct disadvantage of this longitudinal study, which was undertaken to answer the questions that were asked earlier, is that it necessarily includes only those students who attend a Chapter 1 school for three consecutive years, who fall into one of the reading or math program participation sequences, and who had Fall and Spring achievement test scores for three complete years. Another disadvantage is the large population of students necessary to form the groups of participation sequences. The study groups are formulated for the purpose of testing certain hypotheses and should not be construed as necessarily representative of all Chapter 1 students in PPS especially those who were enrolled for less than three years in a Chapter 1 qualifying school.

Subjects. Each year, the number of students in each of grades three through eight enrolled in Chapter 1 ranged from about 850 in grade three to about 500 in grade eight. Although there were usually a few more males (about 55 percent) than females (about 45 percent) enrolled each year, possible sex differences were not considered because the numbers in some groups would have been too small for analyses.

The universe from which study groups were selected included all students in grades three through eight who were in a Chapter 1 school for three consecutive years and who participated in a Chapter 1 reading or math supplementary instruction program for at least one of the three years. The requirement of enrollment in

Chapter 1 schools for the duration of the study period was imposed to preclude student participation in another supplementary program in a non-Chapter 1 school during the year(s) they were not participating in Chapter 1. This requirement was also necessary because qualification for and participation in Chapter 1 was to be used as a basis for predicting performance levels of student groups at the beginning of the three year period. Study groups were defined based on seven possible Chapter 1 participation sequences by years: (1) Y-Y-Y, (2) Y-Y-N, (3) Y-N-Y, (4) N-Y-Y, (5) Y-N-N, (6) N-Y-N, and (7) N-N-Y. Within each participation sequence, there were four possible grade sequences: (1) 3-4-5, (2) 4-5-6, (3) 5-6-7, and (4) 6-7-8.

Study groups were constructed separately for the Chapter 1 reading and math programs, and separately for two different three-year periods: 1983-86 and 1984-87. This methodology was a replication of the earlier design used by Yagi (1984) which included the periods 1979-82 and 1980-83.

Because Fall-to-Spring achievement gain was used as the dependent variable, a further restriction was that students had to have valid Fall and Spring achievement test scores for each year ("clear"), and to have attended the same school from fall to spring within each year ("intact"). Enrollment in the same school for the three year period could not be required because most PPS elementary school students in grade five go to middle schools for grade six.

Independent and Dependent Variables. Participation (Y/N) served as the independent variable and Fall-to-Spring achievement gain as measured by the PALT was the dependent variable. In the 1984 study, the dependent variable was expressed as percent of RIT gain above or below the District average RIT gain. Although percentages of average District gain stir images of amounts that are easily understood, the decision was to use a more refined measure of

growth, a more powerful one statistically. Instead of simply taking the Fall-Spring mean achievement gain for a group and expressing it in standard score units from the District average gain, a correction is applied in recognition of Chapter 1 as a compensatory program for low performing students. Lower than average scores introduce a bias when comparing achievement gains across N and Y years because students during Y years are likely to have lower Fall scores. Despite the equal interval scale and the assignment of test levels according to the student ability levels, there is the possibility of "regression to the mean" in the next testing for extremely low scoring students in the next testing. The bias would tend to exaggerate the achievement gain differences between N and Y that is attributable to the condition purportedly fostered by Chapter 1 intervention. One way chosen here to correct for this initial pre-test difference is through simple regression and the use of a residual gain score.

Standardized achievement growth, i.e., RIT score gains relative to the District average expressed in terms of standard deviations, was regressed on the Fall deviation score (group deviation from the District average) for the entire population of school grade groups in the District. Regressions were performed separately for each study year and for each subject, reading and math. All regressions showed a consistent and statistically significant linear relationship between the Fall score and Fall-to-Spring gain such that the lower the fall scores, the greater the achievement gains.

For each Chapter 1 study group, the residual or distance from the regression line was computed and standardized to indicate how far the obtained score for the group fell above or below the expected or predicted score, scores based on averages. The "standardized residual growth" scores represent achievement gains above or below those which would be predicted by the Fall scores. It controls for pre-test differences and provides a purer measure of growth

attributable to the independent variable, participation or non-participation in Chapter 1.

Predictions/Hypotheses. The gist of the analyses was to compare observed patterns of achievement gains with predicted patterns which were based on assumptions about Chapter 1 and regular school programs and the effect of positive/negative behaviors of students in these programs on achievement test scores. As in the 1984 study, predictions were made without consideration of the participation or non-participation of students in Chapter 1 prior to the first year in the three-year period.

The results were plotted for each study group using bar graphs which showed achievement growth for students as they progressed through three-grade sequences. Bar graphs were used intentionally to replicate those in the original 1984 study, a procedure to facilitate visual inspection. Each graph provides a quick look at the pattern of achievement growth for a particular study group where the three sets of four bars represent participation/non-participation and the first bar in each set is for the grade sequence 3,4,5; the second bar is for grades 4,5,6; and so on for the four grade sequences. The hypotheses and summary of the predicted trend for each study group are as follows:

Groups enrolled in Chapter 1 two or more years

Group	Prediction
Y-Y-Y	Students in this group were expected to have the lowest first year Fall mean score of all groups. This prediction was based on the assumption that students enrolled in Chapter 1 for three successive years were likely to have more problems than others. Due largely to Chapter 1, they were expected to gain about as much as the District average, about equal to the gain for one

subject session, the first year and increase slightly the next two as they adjusted to Chapter 1.

Summary--Graphs (Figure 1) should show random fluctuations around District average with slight increases the last two years.

Y-Y-N Students in this group were expected to have the second lowest Fall mean score. This prediction was based on the assumption much like that for the Y-Y-Y group except their problems were less serious. Gains for this group were expected to be relatively large the first year, perhaps a little larger the second year. Their gains the third year without Chapter 1 would be sustained fairly well, but could drop to equal the District average or slightly above.

Summary--Graphs (Figure 2) should show above average gains the first and second years but drop to about average the third year.

Y-N-Y Students in this group were expected to have the third lowest Fall mean score. This prediction was based on the assumption that their achievement levels were fluctuating around the cutoff. Although their large gains the first year was sufficient to exit them from Chapter 1 the second year, their small gain the second year was sufficient to qualify them again for Chapter 1 the third year. Their gain the third year was expected to be almost as much as the first year, perhaps a trifle less due to persisting problems.

Summary--Graphs (Figure 3) should show a U-shape with above average gains during the program years and less than average in the middle year; slightly smaller gain the third year than the first.

N-Y-Y Among students in Chapter 1 for two or more years, this group was expected to have the highest initial year Fall mean score, but their score would be lower than those for groups in the program for only one year. These were students whose achievements were slipping and recovery likewise was expected to be a little slower. Students in this group were expected to make small gains without Chapter 1 resulting in their being selected for services the second year. Their gains the first year with Chapter 1 were expected to be larger than those without, and gains the next year again with Chapter 1 were expected to be even larger--i.e., larger than the District average.

Summary--Graphs (Figure 4) should show a small gain the first year followed by progressively larger gains the next two years with the third year gain exceeding the District average.

Groups enrolled in Chapter 1 only one year

Group	Prediction
Y-N-N	The initial Fall mean score for this group was expected to be higher than those for groups in Chapter for two or more years but lower than those of groups in Chapter 1 for one year. These were students whose gains were sufficient the first year, perhaps adding to those of previous years, so that services the next two years were unnecessary. Without consistent effort to maintain their positive behavior, their gains would be smaller the second year and perhaps even smaller the third year.

Summary--Graphs (Figure 5) should show a slightly better than District average gain the first year, dropping downward the next two years to below average.

N-Y-N The initial Fall mean score for this group was also expected to be higher than those for groups in Chapter 1 for two or more years and higher than the group in Chapter 1 only the first year. These were students whose achievement scores were near the cut off and a year of supplementary instruction during the second year was sufficient to exit them from Chapter 1. Their gain the third year would be smaller than that for the second year but, hopefully, not as small as that for the first year.

Summary--Graphs (Figure 6) should show a smaller than District average gain the first year, followed by a larger than average gain the second year, returning to about District average gain the third year.

N-N-Y The initial Fall mean score for this group was expected to be the highest of all groups. Their gains the first two years were sufficiently small, smaller than the District average, as to qualify them for Chapter 1 the third year. Their gain the third year would be large, larger than the District average.

Summary--Graphs (Figure 7) should show smaller than District average gains the first two years, followed by a rather large gain up to or slightly exceeding the District average the third year during enrollment in Chapter 1.

Achievement gain patterns were examined visually to determine how closely the predicted patterns matched the obtained. In judging the comparisons, consideration was given to consistency across the four grade sequences and to whether any observed anomalies called the core hypothesis into serious question.

A very simple rating scale of High, Moderate and Low Agreement was used as a means of describing how closely the predicted matched the obtained patterns. The scale is as follows:

High: Very close agreement of predicted and obtained, although one grade sequence may not fit the pattern; hypothesis well-supported.

Medium: Some departures, inconsistencies across grade patterns; mixed support for hypothesis.

Low: Major discrepancies from predicted patterns; hypothesis not well supported.

The decision was made not to combine data from both three-year time periods (1983-86 and 1984-87). Although the general patterns of gains for corresponding groups were alike, some of the data points were sufficiently discrepant that combining data across the two time periods is potentially misleading. Besides, replication over two time periods within the current study was viewed as an advantage. Combining the two time periods and increasing the sample sizes would have resulted in greater statistical power, but this was not a concern because the analysis did not involve statistical significance testing.

RESULTS

Group performances for both reading and math were predicted to be essentially the same; however, results for each subject will be presented separately to enable those interested in subject matter differences to study the findings. Judgments by the authors of how closely the predicted patterns matched the obtained are presented first (More details of all predictions, brief summaries of obtained results, and judgments of how closely the predicted matched the obtained results are in Appendix A.). These are followed by sets of bar graphs, Figures 1 to 7, which show gains expressed as standardized residual growth for reading and math for each three-year time interval (Note: Bar graph values plotted at +3 to -3 are three standard deviations or more from the mean; exact values are shown in the tables.). The graphs are followed by sets of corresponding tables, Tables 1-4, that present the same results numerically including sample sizes. A final table, Table 5, shows the first year Fall mean scores for each group arranged in their predicted order.

Following is a summary of the ratings of the extent of agreement between predicted and obtained results for each group in reading and math for the two three-year periods:

		Agreement of Predicted and Obtained	
GROUP	SUBJECT	1983-86	1984-87
YYY	Reading	High	High
	Math	Moderate	Moderate
YYN	Reading	Moderate	Low
	Math	Low	Moderate
YNY	Reading	High	High
	Math	Moderate	Moderate
NYY	Reading	Moderate	High
	Math	High	Moderate
YNN	Reading	High	Moderate
	Math	High	Moderate

NYN	Reading	Moderate	Moderate
	Math	Low	Low
NNY	Reading	High	High
	Math	High	High

Summary of Findings--The findings are summarized in terms of (1) two time periods, (2) four grade sequences, (3) two subjects, and (4) predicted and obtained patterns of gains.

1. Within the current study, there was considerable consistency in the results across the two time periods, 1983-86 and 1984-87. The results tend to replicate findings by Yagi (1984) for the combined periods 1979-82 and 1980-1983. The statements are equally true for reading and math.

2. Although results for all grade sequences were predicted quite well, the predictions tended to be less accurate for sequences spanning the middle school years. Perhaps experiences in transferring from self-contained elementary school classes to separately scheduled classes in middle schools might have rendered group achievement scores more unpredictable. These findings do not necessarily reflect a lack of quality or consistency in upper grade Chapter 1 programs.

3. Predicted results for reading were closer to the obtained than those for math. Of the 14 predictions for reading, 8 were rated high agreement, 5 moderate, and 1 low agreement. Of the 14 predictions for math, 4 were rated high agreement, 7 moderate, and 3 low agreement.

4. The predicted Fall scores taken as starting points for the groups were predicted quite well especially for all reading groups and for the extreme math groups, i.e.,

lowest scores for Y-Y-Y and Y-Y-N and the highest for N-Y-N and N-N-Y. The results demonstrated the predicted increases in achievement gains while students were in Chapter 1 quite well, but their gains were not sustained as well as expected the first year following Chapter 1. Sustained gains tended to be over estimated especially for the Y-Y-N and N-Y-N groups.

DISCUSSION

The longitudinal study of students in and out of a Chapter 1 program over a three year period was undertaken to answer four critical questions: (1) Are Chapter 1 programs working, i.e., producing noticeable achievement gains? (2) What works and what doesn't? (3) Why are some students in the program year after year? and (4) Do students sustain their gains? Answers to these questions rested heavily on the answer to the most significant question,"Are patterns of student achievement over the three year period predictable?"

The present study and the previous study by Yagi (1984) were both based on the premise that students would make greater achievement gains while they were in Chapter 1 reading or math programs than when they were not--a period of time when positive behaviors was maximized. With students serving as their own controls when they were in and out of Chapter 1, and with predicted achievement gain patterns matching the obtained quite closely, the data strongly support the contention that, as a group, PPS Chapter 1 programs are working. The data also support the contention that although some degree of quality is necessary, the key feature of Chapter 1 programs was not so much the curricular quality of these programs but the extent students engaged in socially acceptable and educationally worthwhile behaviors. Smaller achievement gains,

often noticeably less than the District average, were predicted when these students were not in Chapter 1 and negative behaviors predominated. The fact that score patterns were predicted quite accurately for both subjects, even in terms of amounts for different groups, provides strong evidence to support the contention the combination of socially acceptable and educationally worthwhile behavior is the most significant of all effective school practices to be identified to date. These positive behaviors are not just limited to time during seat work or learning tasks, but it includes all time in the classroom; between classes; and during work, play, and meals. The behaviors are characteristic of being a more responsible person as well as learner. Thus, single measures, for example, of on-task behavior during a defined period or periods of observation fall far short of being adequate in measuring socially acceptable and educationally worthwhile behaviors.

Why are some students in Chapter 1 year after year? In PPS, there are Chapter 1 programs for reading and math, and in any given year a student may be in either one or both so a student may move from one to the other. Consider also that students in the Y-Y-Y and Y-Y-N groups had the lowest Fall scores during the first year of study. They had the most problems and largest deficit to overcome. The predicted Fall scores were based on this hypothesis.

An answer to the question about sustained gains is "it depends." As predicted, students sustained their gains following Chapter 1 if they were again in Chapter 1; they did not, at least not as well, if they were not. The prediction that students in the Y-Y-N group would sustain their gains quite well was based on the premise that positive behaviors in Chapter 1 for two years would carry over to the third year. The extent of the carryover was noticeably overestimated because students apparently tended to revert to their original behavior. Carryover after only one year

in Chapter 1, e.g., N-Y-N group, was expected to be minimal. This is primarily because increased conscientious efforts to try to sustain or increase the positive behaviors of these students in the regular classroom and elsewhere were not expected to be forthcoming.

Many Chapter 1 research studies, especially sustained gain studies, have produced mixed or even contradictory results mainly because they failed to take into account the experiences of students in their regular classrooms. In this study, it was assumed that student achievement scores were affected by their total school experiences; and, obviously, students who engaged in positive behaviors were expected to make larger gains than those who did not. It should also be obvious that the extent students engage in positive behaviors varies from student to student and from school to school for both regular classrooms and Chapter 1 programs; therefore, one could not predict results for a particular school without knowing something about all three, i.e., students, regular classroom, Chapter 1 program. Although it was acknowledged that differences due to teachers, students (e.g., SES, sex, ethnicity), subjects, program quality, etc., either alone or in combination could be more or less significant for a particular study group, for the sake of this study it was generally assumed that these were equal for all groups. On the other hand, the extent of positive behavior, especially in regular classrooms, was not assumed to be equal for all groups.

The fact that performances in math were not predicted quite as well as performances in reading is puzzling and the investigators are left without an explanation. Differences due to incidental learning opportunities, nature of the subjects, etc., may be involved. Even though there is a temptation to figure out how to improve upon the predictions, the intention here is not to but to propose how the theory of socially acceptable and educationally

worthwhile behaviors might be tested in practice. It is anticipated that when attention is focused specifically on socially acceptable and educationally worthwhile behavior, student performances in total, not only in reading and math, will improve. In other words, we are confident that when those who deal with Chapter 1 children in regular classrooms know exactly what needs to be accomplished, they can and will figure out ways to accomplish just that.

If one agrees with the theory and the conclusions from this study, there are implications for research studies and performance improvement to take place simultaneously. For example, what happens when the Chapter 1 effort is strongly supported by the regular classroom and entire school where positive behaviors of these students are maximized? Our predictions were based on the premise that the experiences are additive; and, we would predict that with pull-out programs, the gains would be larger than the expected and could even double! What happens with other combinations where behaviors in Chapter 1 or regular classrooms are not well managed? What are the implications for pull-out vs in-class Chapter 1 programs, a long standing argument? Over the years, PPS Chapter 1 programs have been almost exclusively pull-out, and Chapter 1 gains for some schools are noticeably larger than the gains for the school as a whole even though the gains of the latter are average or above. Obtaining gains that are noticeably larger than the expected could be difficult with in-class instruction--the additional instructional time could to make a difference. How much can performances of Chapter 1 students be improved anyway? There is little evidence to indicate that Chapter 1 students are any less capable than the average students. PPS Chapter 1 programs now participating in program improvement appear to be recognizing the contributions of the entire school to Chapter 1 student performances and the need for close coordination and cooperation. Perhaps consistent conscientious efforts will be made

to increase socially acceptable and educationally worthwhile behaviors of Chapter 1 students. Deliberate and successful attempts to manipulate positive behaviors are necessary to test the theory.

FIGURE 1 -READING. YES-YES-YES

RESIDUAL GROWTH SCORES 83-84, 84-85, 85-86

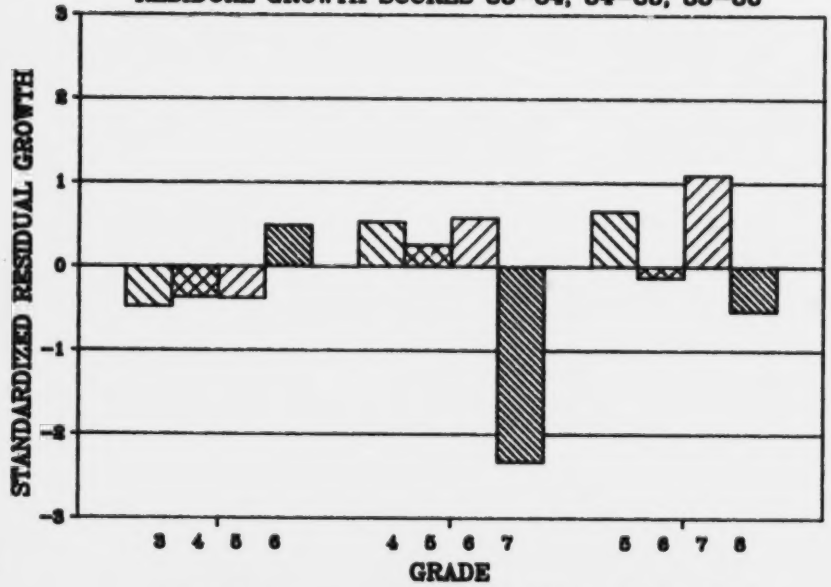
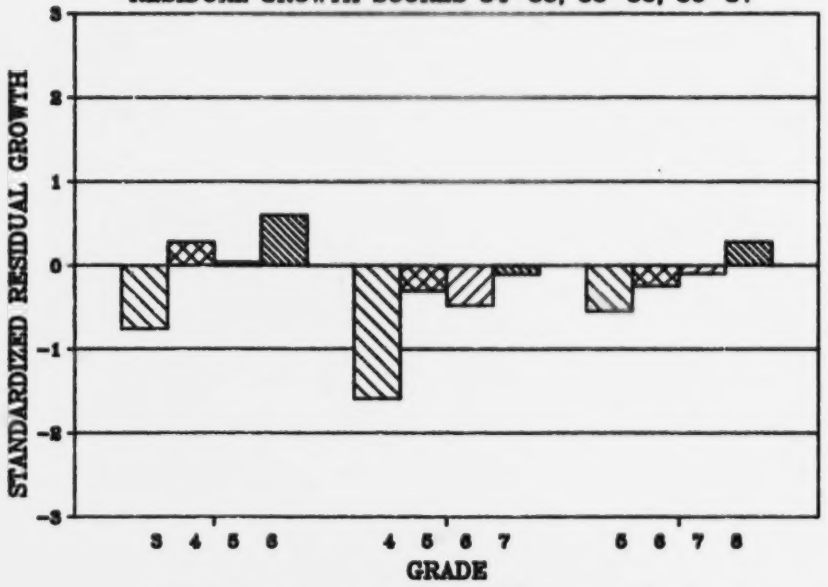


FIGURE 1 -READING. YES-YES-YES

RESIDUAL GROWTH SCORES 84-85, 85-86, 86-87



30

FIGURE 1 -MATH. YES-YES-YES

RESIDUAL GROWTH SCORES 83-84, 84-85, 85-86

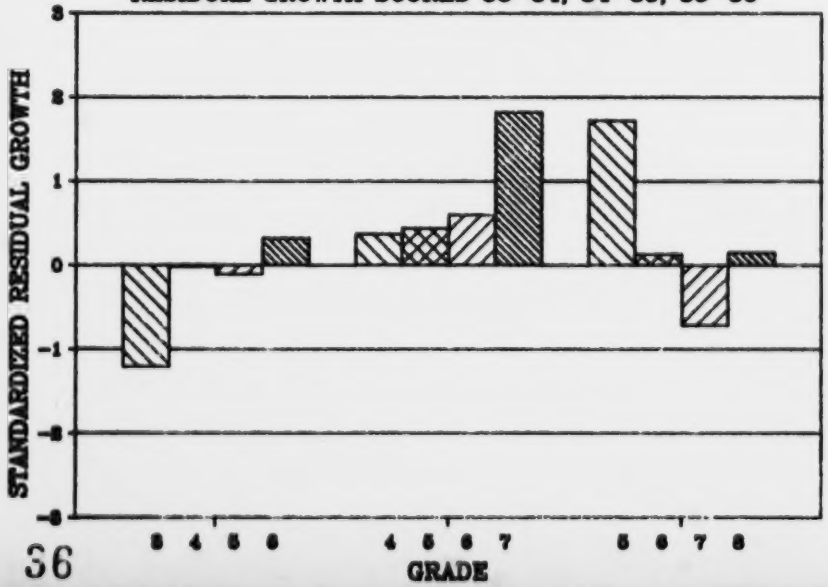


FIGURE 1 -MATH. YES-YES-YES

RESIDUAL GROWTH SCORES 84-85, 85-86, 86-87

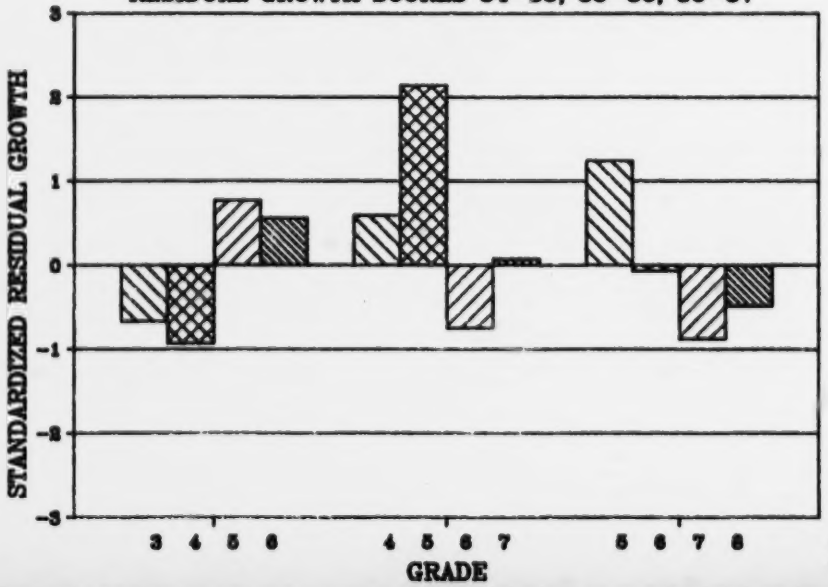


FIGURE 2 -READING. YES-YES-NO

RESIDUAL GROWTH SCORES 83-84, 84-85, 85-86

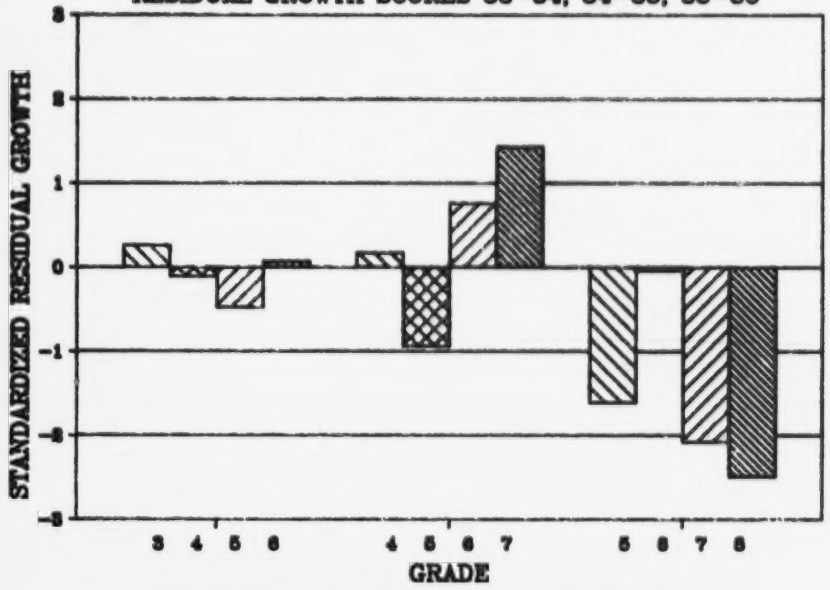


FIGURE 2 -READING. YES-YES-NO

RESIDUAL GROWTH SCORES 84-85, 85-86, 86-87

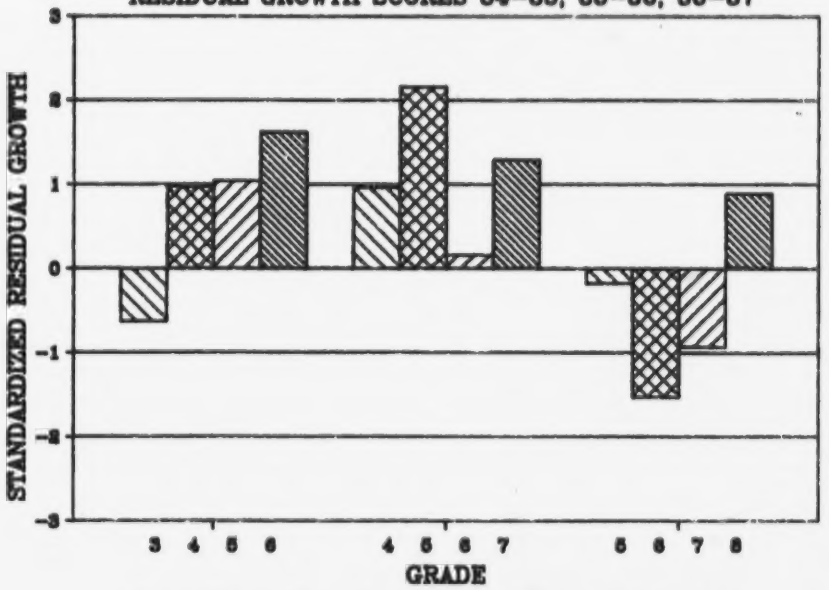


FIGURE 2 -MATH. YES-YES-NO

RESIDUAL GROWTH SCORES 83-84, 84-85, 85-86

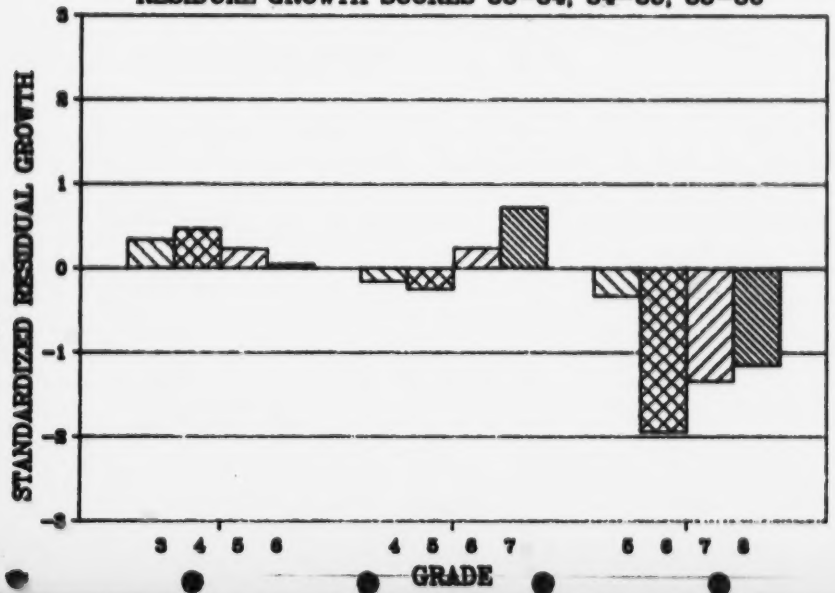


FIGURE 2 -MATH. YES-YES-NO

RESIDUAL GROWTH SCORES 84-85, 85-86, 86-87

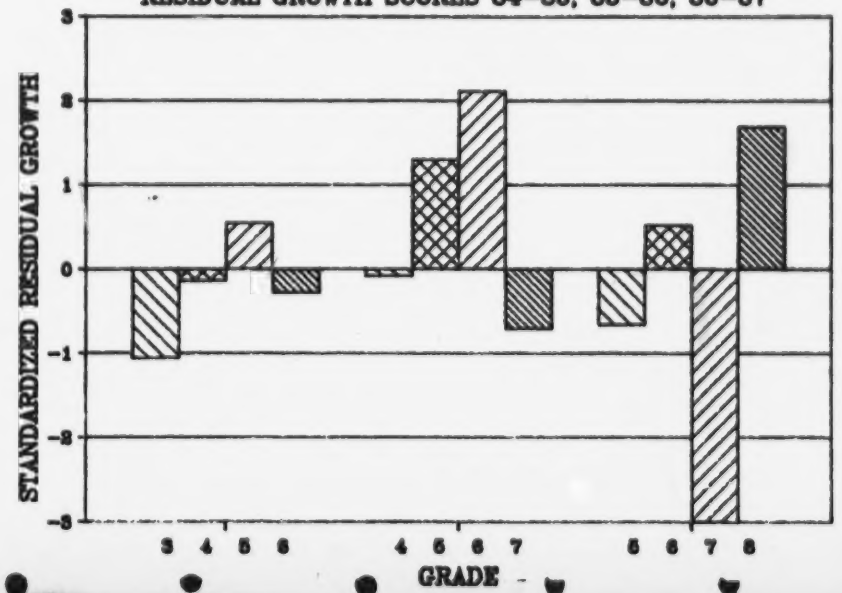


FIGURE 3 –READING. YES–NO–YES

RESIDUAL GROWTH SCORES 83–84, 84–85, 85–86

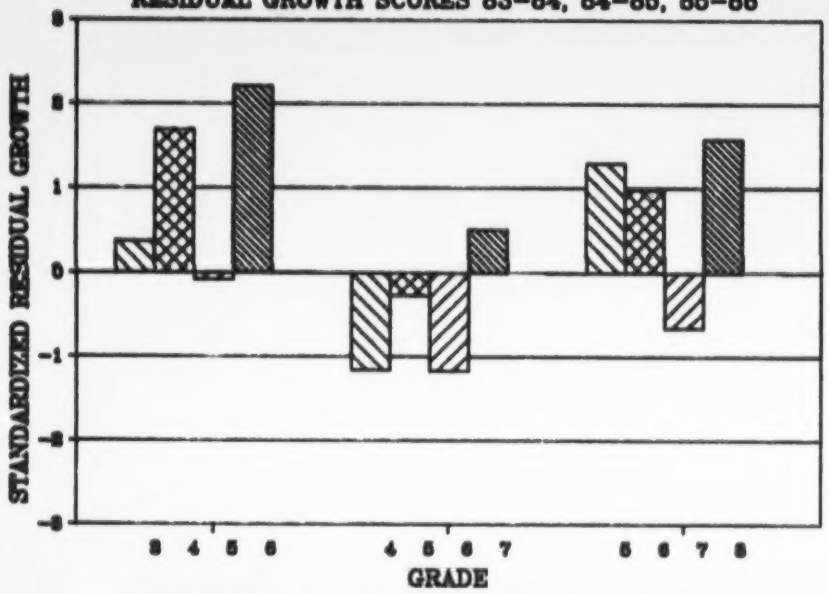


FIGURE 3 –READING. YES–NO–YES

RESIDUAL GROWTH SCORES 84–85, 85–86, 86–87

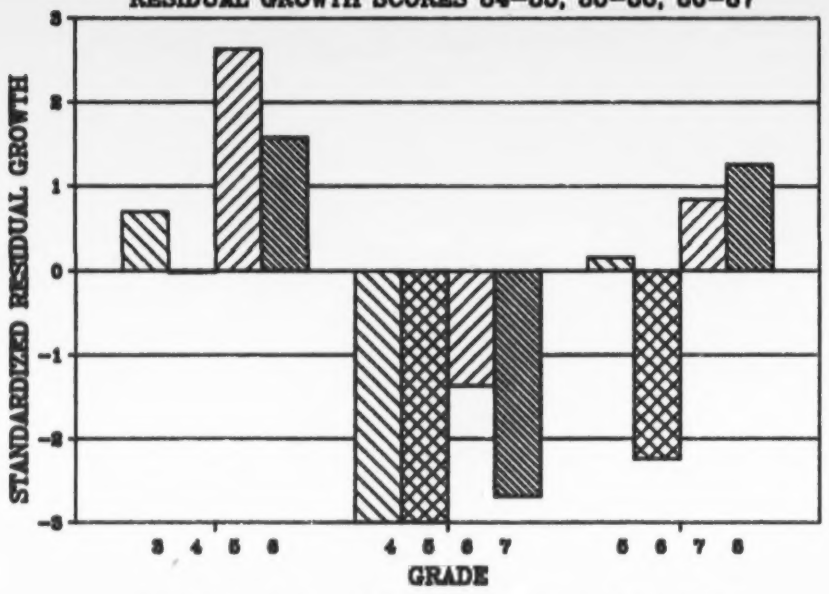


FIGURE 3 –MATH. YES–NO–YES

RESIDUAL GROWTH SCORES 83–84, 84–85, 85–86

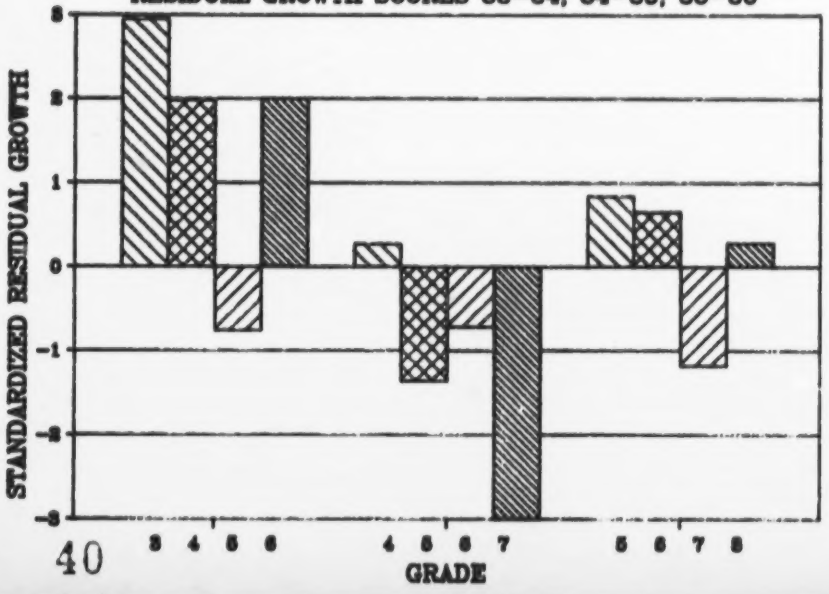


FIGURE 3 –MATH. YES–NO–YES

RESIDUAL GROWTH SCORES 84–85, 85–86, 86–87

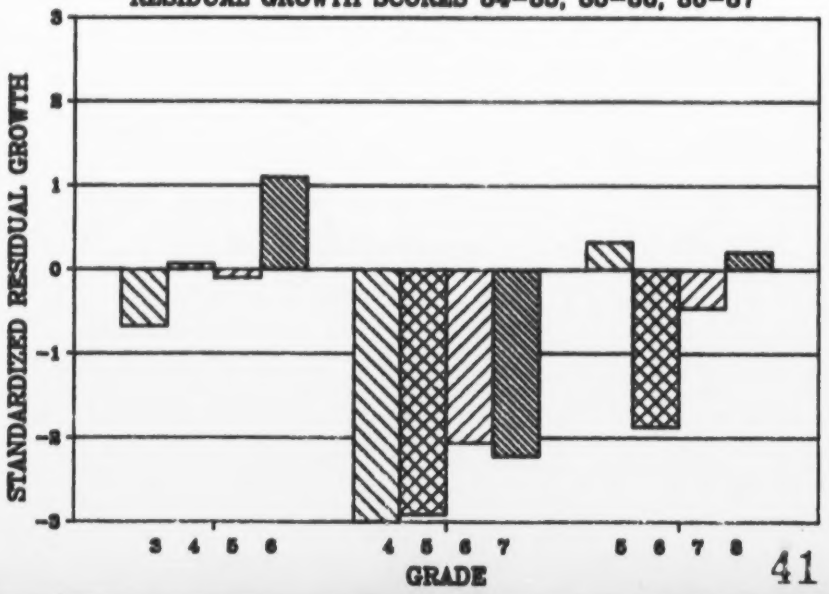


FIGURE 4 -READING. NO-YES-YES

RESIDUAL GROWTH SCORES 83-84, 84-85, 85-86

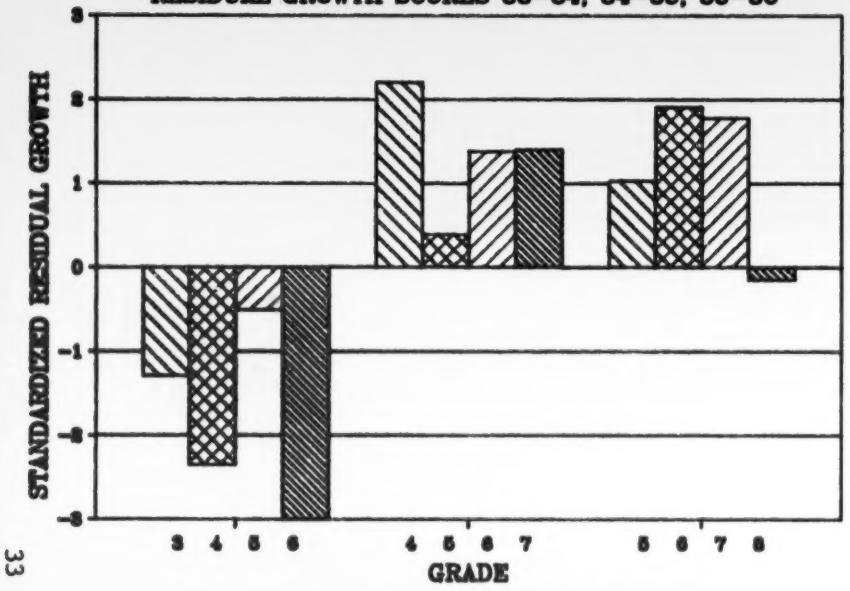


FIGURE 4 -READING. NO-YES-YES

RESIDUAL GROWTH SCORES 84-85, 85-86, 86-87

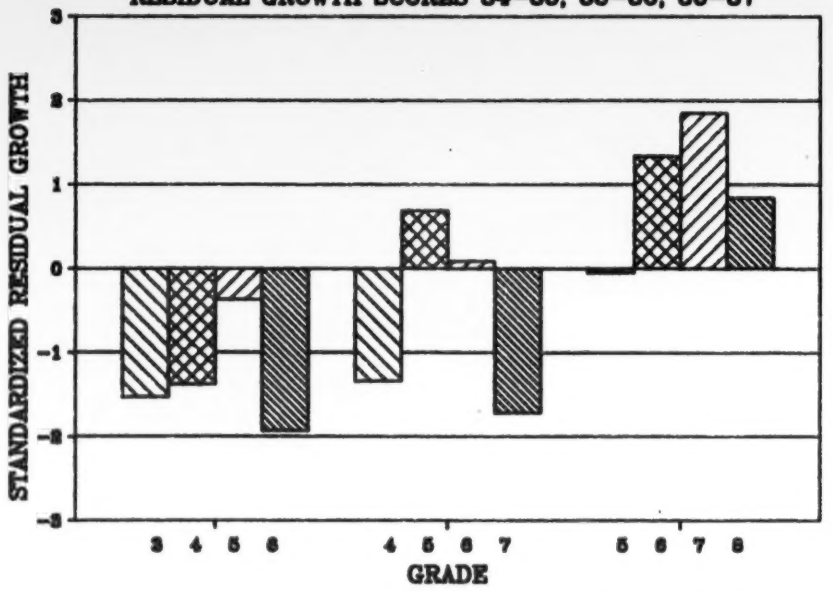


FIGURE 4 -MATH. NO-YES-YES

RESIDUAL GROWTH SCORES 84-85, 85-86, 86-87

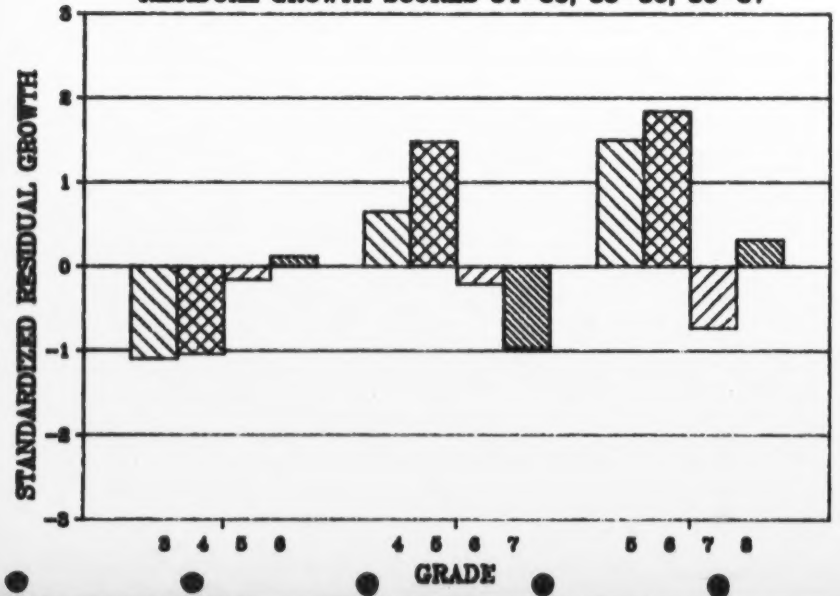


FIGURE 4 -MATH. NO-YES-YES

RESIDUAL GROWTH SCORES 84-85, 85-86, 86-87

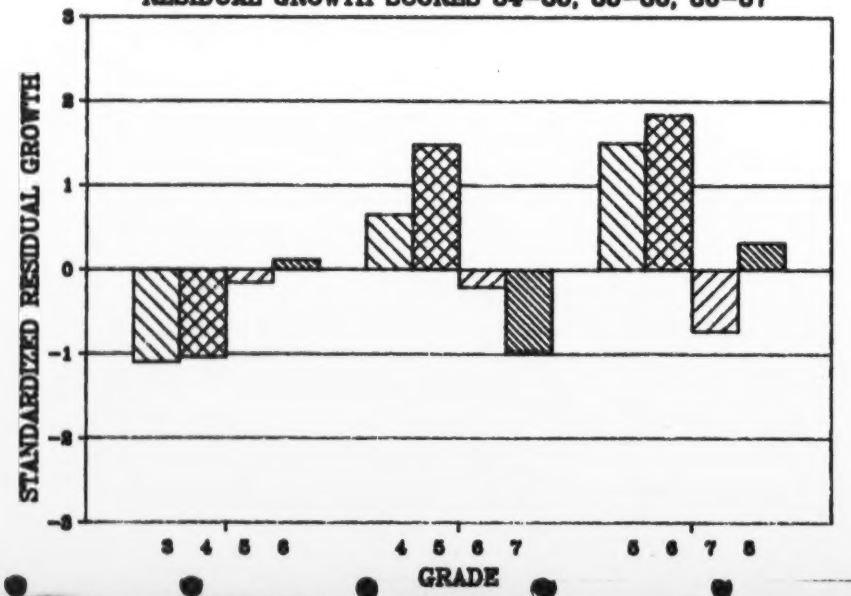


FIGURE 5 -READING. YES-NO-NO

RESIDUAL GROWTH SCORES 83-84, 84-85, 85-86

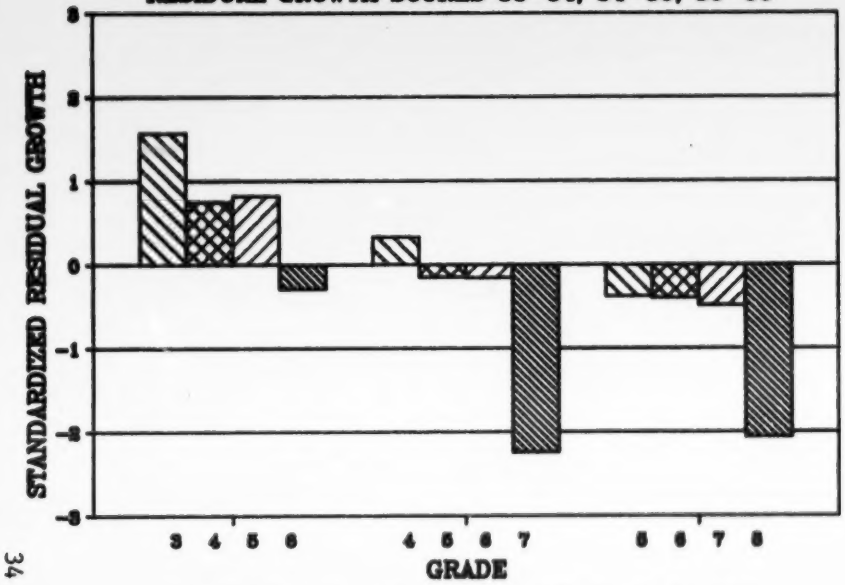


FIGURE 5 -READING. YES-NO-NO

RESIDUAL GROWTH SCORES 84-85, 85-86, 86-87

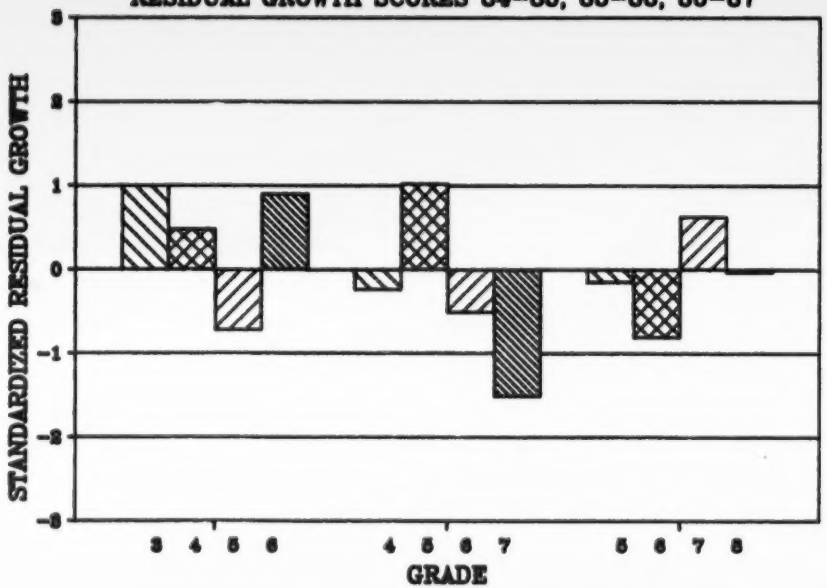


FIGURE 5 -MATH. YES-NO-NO

RESIDUAL GROWTH SCORES 83-84, 84-85, 85-86

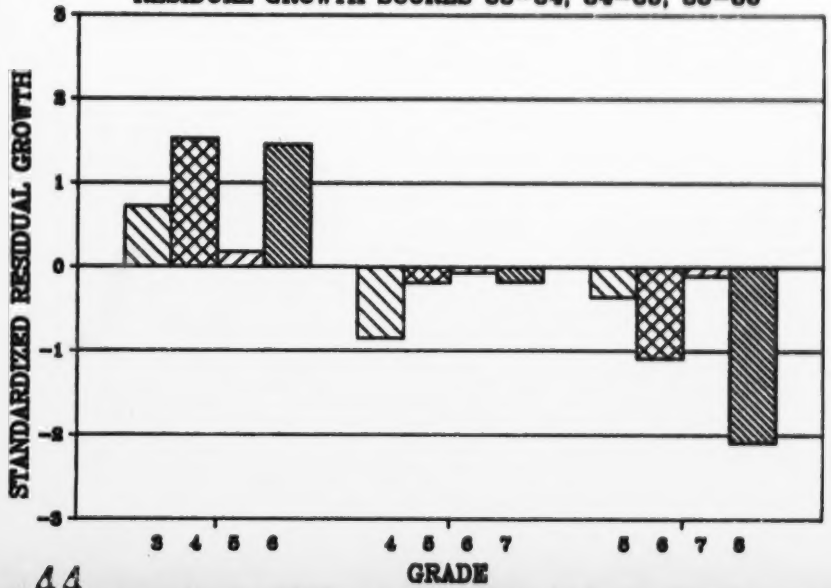


FIGURE 5 -MATH. YES-NO-NO

RESIDUAL GROWTH SCORES 84-85, 85-86, 86-87

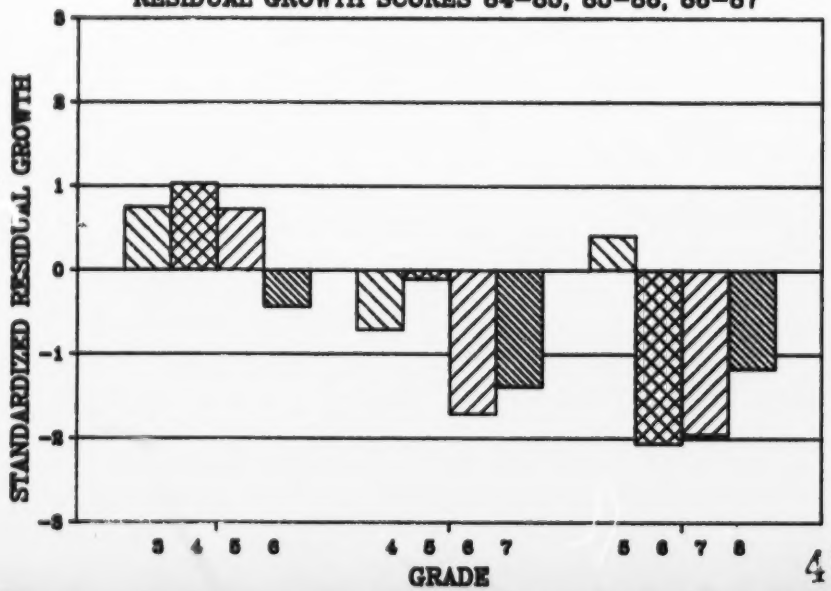


FIGURE 6 -READING. NO-YES-NO

RESIDUAL GROWTH SCORES 83-84, 84-85, 85-86

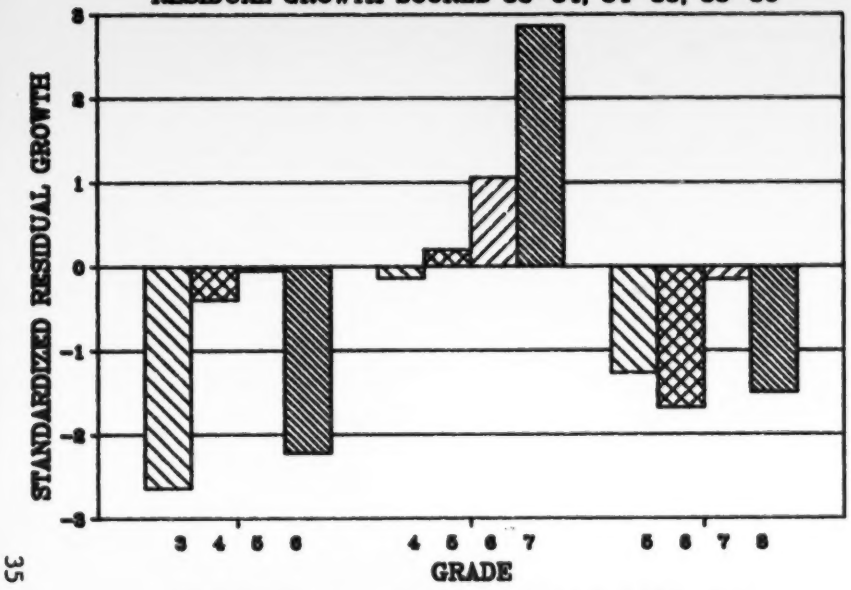


FIGURE 6 -READING. NO-YES-NO

RESIDUAL GROWTH SCORES 84-85, 85-86, 86-87

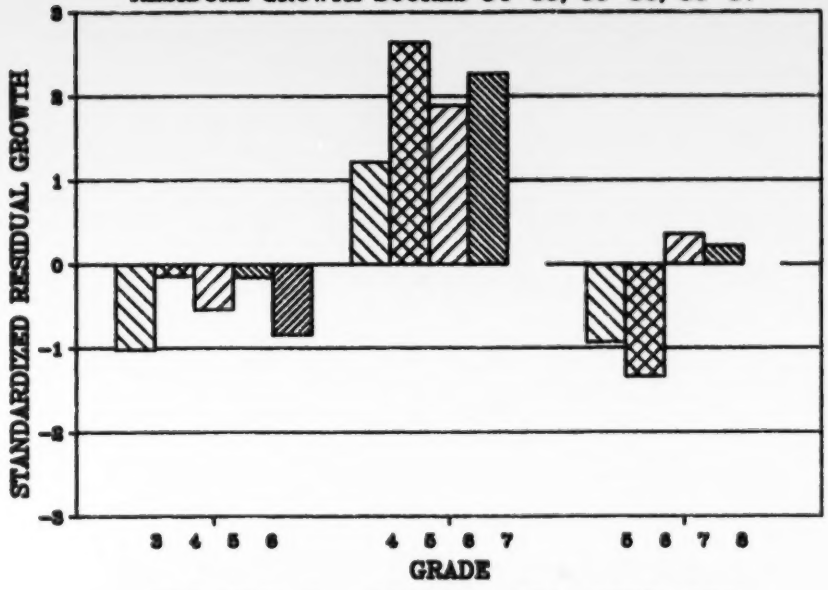


FIGURE 6 -MATH. NO-YES-NO

RESIDUAL GROWTH SCORES 83-84, 84-85, 85-86

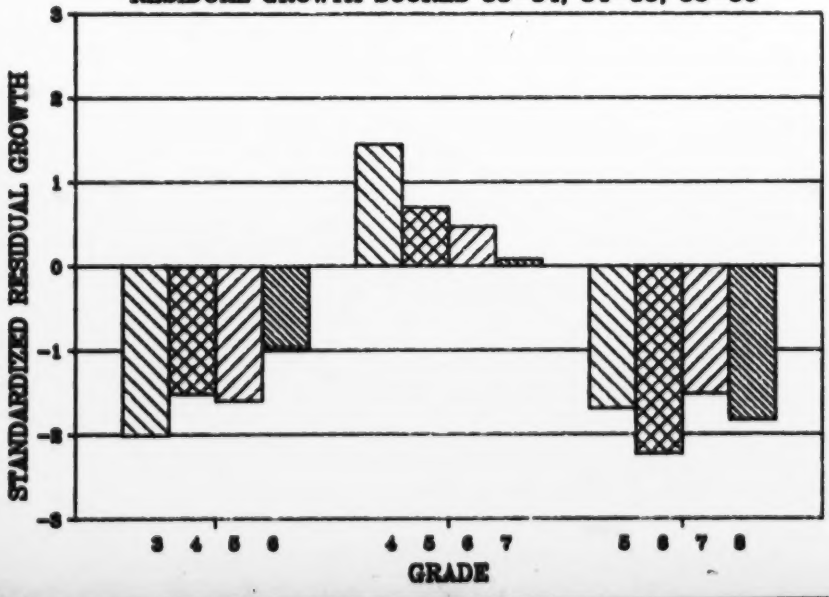


FIGURE 6 -MATH. NO-YES-NO

RESIDUAL GROWTH SCORES 84-85, 85-86, 86-87

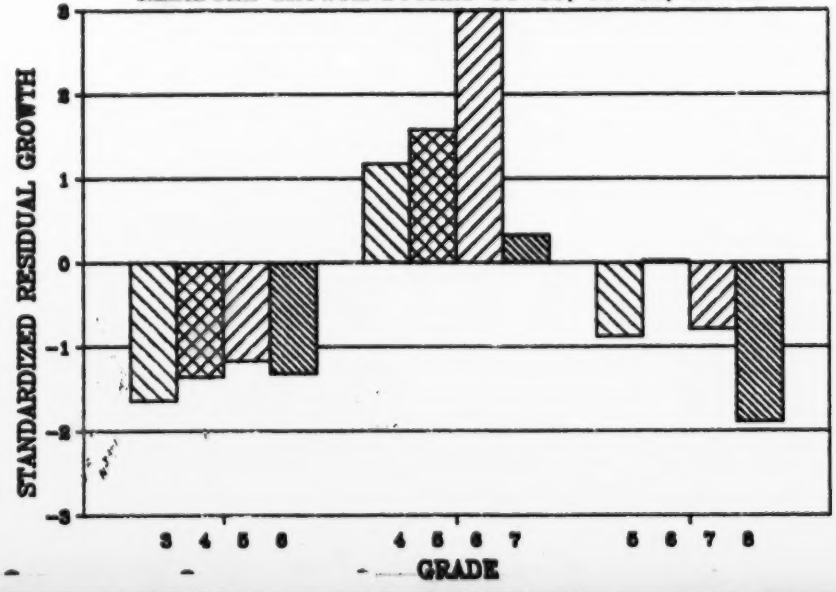


FIGURE 7 -READING. NO-NO-YES

RESIDUAL GROWTH SCORES 83-84, 84-85, 85-86

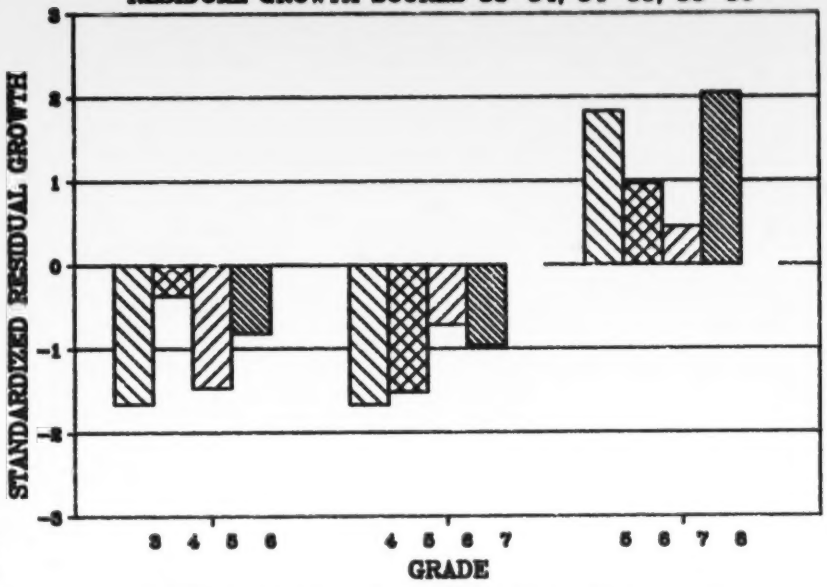


FIGURE 7 -READING. NO-NO-YES

RESIDUAL GROWTH SCORES 84-85, 85-86, 86-87

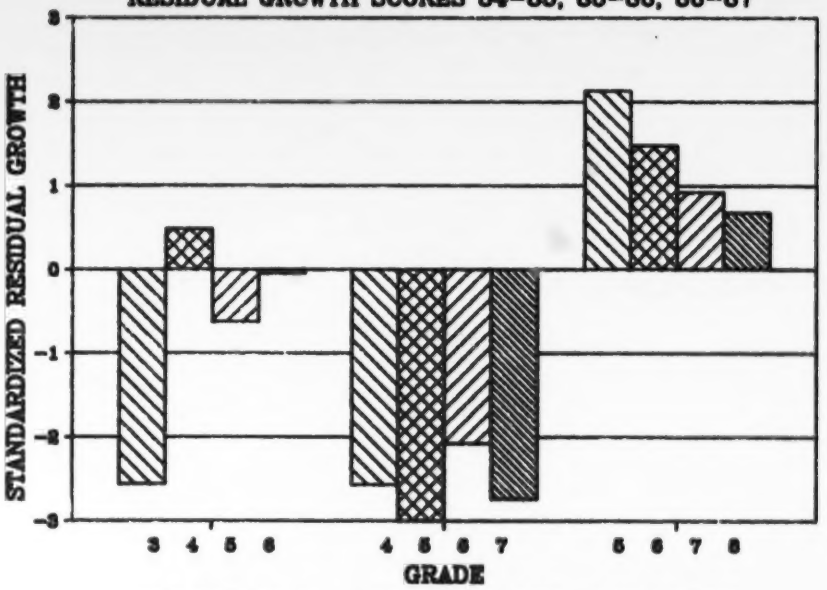


FIGURE 7 -MATH. NO-NO-YES

RESIDUAL GROWTH SCORES 83-84, 84-85, 85-86

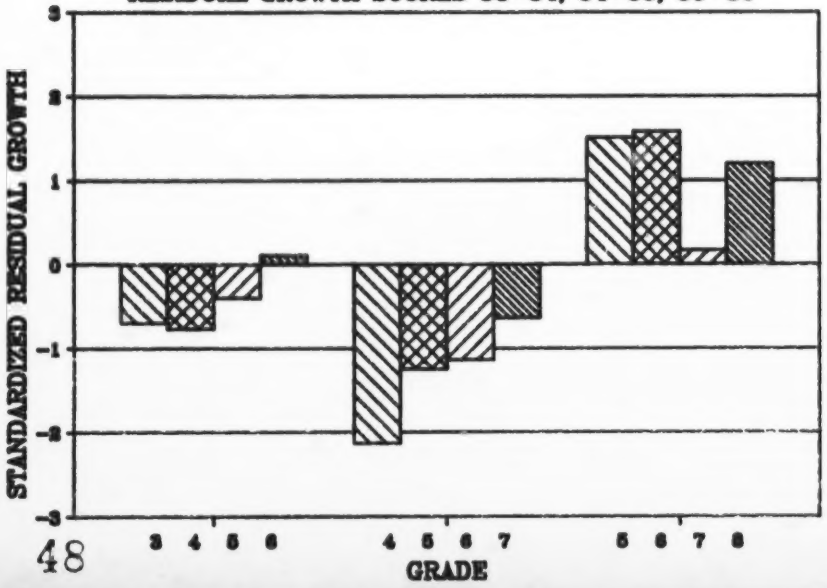


FIGURE 7 -MATH. NO-NO-YES

RESIDUAL GROWTH SCORES 84-85, 85-86, 86-87

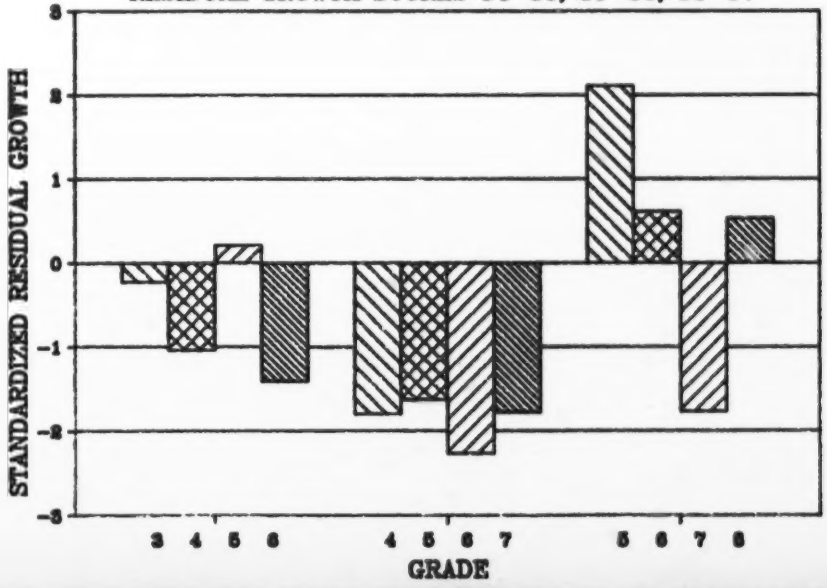


TABLE 1.

1983-86 ANALYSIS: READING DEVIATION, GROWTH, STD, RESIDUAL GROWTH.

RGROUP	GRADE	FALL DEV	SPR DEV	GROW	SRESID	N	YR-SUBJ
1. Y-Y-Y	3	-4.06	-4.00	.16	-.48	103	34R
	4	-3.72	-3.30	1.22	.53	103	45R
	5	-3.53	-3.19	1.07	.66	103	56R
	4	-3.42	-3.36	.17	-.37	83	34R
	5	-3.01	-2.78	.82	.26	83	45R
	6	-2.95	-2.88	.30	-.13	83	56R
	5	-3.07	-3.04	.11	-.38	63	34R
	6	-3.40	-3.13	1.22	.58	63	45R
	7	-3.64	-3.21	1.46	1.09	63	56R
	6	-3.33	-3.09	.99	.49	58	34R
	7	-2.73	-3.01	-1.93	-2.33	58	45R
	8	-2.99	-3.01	-.04	-.53	58	56R
2. Y-Y-N	3	-3.02	-2.71	.72	.26	74	34R
	4	-2.20	-2.01	.59	.17	74	45R
	5	-1.80	-2.18	-1.14	-1.61	74	56R
	4	-2.67	-2.57	.31	-.11	42	34R
	5	-2.48	-2.61	-.52	-.94	42	45R
	6	-2.23	-2.16	.29	-.04	42	56R
	5	-2.56	-2.58	-.07	-.48	47	34R
	6	-2.54	-2.25	1.26	.76	47	45R
	7	-2.29	-2.66	-1.49	-2.08	47	56R
	6	-2.65	-2.53	.48	.07	53	34R
	7	-2.48	-2.21	1.96	1.43	53	45R
	8	-1.94	-2.34	-1.89	-2.49	53	56R
3. Y-N-Y	3	-3.55	-3.17	.90	.36	51	34R
	4	-2.53	-2.83	-.74	-1.16	51	45R
	5	-3.18	-2.68	1.57	1.29	51	56R
	4	-2.86	-2.18	2.08	1.69	34	34R
	5	-1.78	-1.75	.06	-.27	34	45R
	6	-2.24	-1.84	1.18	.98	34	56R
	5	-1.99	-1.92	.23	-.08	29	34R
	6	-2.04	-2.27	-.84	-1.17	29	45R
	7	-2.59	-2.61	-.21	-.66	29	56R
	6	-1.88	-1.28	2.43	2.21	13	34R
	7	-1.45	-1.34	.81	.50	13	45R
	8	-1.51	-1.20	1.61	1.58	13	56R
4. N-Y-Y	3	-2.08	-2.50	-.93	-1.29	18	34R
	4	-2.76	-1.75	2.81	2.20	18	45R
	5	-1.99	-1.62	1.19	1.03	18	56R
	4	-1.56	-2.21	-2.03	-2.35	19	34R
	5	-2.54	-2.29	.88	.39	19	45R
	6	-2.18	-1.48	1.99	1.91	19	56R
	5	-2.87	-2.88	-.05	-.51	27	34R
	6	-3.13	-2.66	2.01	1.38	27	45R
	7	-3.18	-2.61	2.00	1.78	27	56R
	6	-2.59	-3.28	-2.78	-3.28	20	34R
	7	-2.86	-2.58	1.99	1.40	20	45R
	8	-2.51	-2.47	.23	-.15	20	56R

TABLE 2.

1983-86 ANALYSIS: MATH DEVIATION, GROWTH, STD, RESIDUAL GROWTH.

RGROUP	GRADE	FALL DEV	SPR DEV	GROW	SRESID	N	YR-SUBJ
1. Y-Y-Y	3	-3.88	-4.10	-.41	-1.21	69	34M
	4	-3.65	-3.20	1.32	.37	69	45M
	5	-3.60	-2.80	2.20	1.72	69	56M
	4	-3.47	-3.25	.53	-.02	70	34M
	5	-3.41	-2.91	1.35	.44	70	45M
	6	-3.08	-2.91	.64	.13	70	56M
	5	-3.30	-2.92	.42	-.11	65	34M
	6	-3.29	-3.02	1.50	.60	65	45M
	7	-2.88	-2.90	-.18	-.72	65	56M
	6	-3.08	-2.84	.75	.32	64	34M
	7	-3.02	-2.58	2.79	1.82	64	45M
	8	-2.83	-2.69	.61	.15	64	56M
2. Y-Y-N	3	-2.66	-2.30	.69	.34	62	34M
	4	-2.19	-2.08	.40	-.16	62	45M
	5	-2.04	-2.05	.03	-.33	62	56M
	4	-3.00	-2.63	.86	.47	39	34M
	5	-2.40	-2.28	.35	-.25	39	45M
	6	-2.39	-2.77	-1.40	-1.94	39	56M
	5	-3.57	-2.94	.75	.23	43	34M
	6	-3.53	-3.33	1.15	.24	43	45M
	7	-3.33	-3.50	-.68	-1.34	43	56M
	6	-2.50	-2.36	.42	.05	40	34M
	7	-2.57	-2.35	1.47	.72	40	45M
	8	-2.37	-2.49	-.67	-1.15	40	56M
3. Y-N-Y	3	-2.63	-1.12	2.89	2.95	32	34M
	4	-1.04	-.85	.62	.27	32	45M
	5	-1.28	-.93	.98	.84	32	56M
	4	-2.40	-1.53	2.03	1.98	27	34M
	5	-1.72	-2.13	-1.04	-1.36	27	45M
	6	-2.05	-1.08	.94	.65	27	56M
	5	-1.39	-1.65	-.45	-.76	23	34M
	6	-1.83	-1.96	-.30	-.72	23	45M
	7	-2.04	-2.23	-.75	-1.18	23	56M
	6	-2.15	-1.53	2.01	2.00	18	34M
	7	-1.62	-2.11	-2.89	-3.01	18	45M
	8	-2.08	-1.94	.60	.28	18	56M
4. N-Y-Y	3	-3.34	-3.90	-1.05	-1.86	38	34M
	4	-3.63	-3.60	.18	-.66	38	45M
	5	-3.84	-2.99	2.33	1.82	38	56M
	4	-2.95	-2.96	-.02	-.56	19	34M
	5	-2.77	-2.91	-.34	-.95	19	45M
	6	-2.59	-2.27	1.22	.85	19	45M
	5	-2.55	-3.01	-.72	-1.31	25	34M
	6	-2.78	-2.84	.02	-.63	25	45M
	7	-3.06	-2.70	1.12	.66	25	56M
	6	-2.78	-3.15	-1.29	-2.03	29	34M
	7	-2.71	-2.84	-.69	-1.25	29	45M
	8	-2.69	-2.49	.93	.52	29	56M

TABLE 3.

1984-87 ANALYSIS: READING DEVIATION, GROWTH, STD, RESIDUAL GROWTH.

RGROUP	GRADE	FALL DEV	SPR DEV	GROW	SRESID	N	YR-SUBJ
1. Y-Y-Y	3	-3.87	-3.92	-.10	-.76	74	45R
	4	-3.45	-3.73	-.90	-1.59	74	56R
	5	-3.52	-3.51	-.07	-.55	74	67R
	4	-4.09	-3.74	1.02	.28	76	45R
	5	-3.84	-3.77	.26	-.31	76	56R
	6	-3.59	-3.56	.27	-.25	76	67R
	5	-3.20	-3.02	.63	.04	64	45R
	6	-3.13	-3.16	.02	-.48	64	56R
	7	-2.92	-2.85	.33	-.10	64	67R
	6	-3.42	-3.14	1.24	.60	61	45R
	7	-3.46	-3.32	.39	-.11	61	56R
	8	-2.90	-2.81	.75	.28	61	67R
2. Y-Y-N	3	-2.75	-2.82	-.16	-.63	74	45R
	4	-2.33	-1.97	1.18	.96	74	56R
	5	-2.33	-2.26	.15	-.18	74	67R
	4	-3.30	-2.73	1.62	.97	49	45R
	5	-2.85	-2.12	2.30	2.16	49	56R
	6	-2.02	-2.28	-1.39	-1.53	49	67R
	5	-3.00	-2.54	1.64	1.04	46	45R
	6	-2.42	-2.28	.49	.16	46	56R
	7	-2.19	-2.28	-.70	-.93	46	67R
	6	-3.34	-2.80	2.30	1.62	33	45R
	7	-3.10	-2.65	1.56	1.29	33	56R
	8	-2.85	-2.67	1.41	.89	33	67R
3. Y-N-Y	3	-3.02	-2.48	1.27	.69	17	45R
	4	-1.74	-2.53	-2.54	-3.21	17	56R
	5	-2.70	-2.50	.59	.16	17	67R
	4	-2.84	-2.68	.50	-.02	33	45R
	5	-2.00	-2.90	-2.77	-3.51	33	56R
	6	-2.47	-2.86	-2.11	-2.24	33	67R
	5	-3.11	-2.19	3.32	2.63	13	45R
	6	-2.72	-3.05	-.80	-1.37	13	56R
	7	-2.51	-2.28	1.31	.84	13	67R
	6	-3.43	-2.89	2.28	1.58	23	45R
	7	-3.29	-3.76	-1.88	-2.69	23	56R
	8	-2.93	-2.69	1.84	1.26	23	67R
4. N-Y-Y	3	-2.56	-3.05	-1.12	-1.53	39	45R
	4	-3.06	-3.29	-.74	-1.34	39	56R
	5	-3.00	-2.86	.40	-.05	39	67R
	4	-2.61	-2.99	-.96	-1.38	39	45R
	5	-3.80	-3.45	1.13	.69	39	56R
	6	-2.94	-2.61	1.93	1.34	39	67R
	5	-2.46	-2.43	.07	-.37	34	45R
	6	-2.50	-2.37	.44	.09	34	56R
	7	-2.19	-1.79	2.37	1.85	34	67R
	6	-2.96	-3.34	-1.48	-1.93	34	45R
	7	-3.73	-3.96	-.98	-1.72	34	56R
	8	-2.83	-2.65	1.35	.84	34	67R

TABLE 4.

1984-87 ANALYSIS: MATH DEVIATION, GROWTH, STD, RESIDUAL GROWTH.

RGROUP	GRADE	FALL DEV	SPR DEV	GROW	SRESID	N	YR-SUBJ
1. Y-Y-Y	3	-4.00	-3.85	.25	-.67	40	45M
	4	-3.84	-3.31	1.19	.59	40	56M
	5	-3.65	-3.11	1.96	1.24	40	67M
	4	-3.93	-3.99	-.05	-.93	43	45M
	5	-4.84	-3.82	2.80	2.14	43	56M
	6	-3.97	-3.86	.64	-.07	43	67M
	5	-3.76	-3.09	1.79	.77	56	45M
	6	-3.11	-3.16	-.17	-.75	56	56M
	7	-2.94	-3.02	-.39	-.88	56	67M
	6	-3.42	-3.15	1.48	.56	58	45M
	7	-3.23	-3.02	.60	.07	58	56M
	8	-2.71	-2.72	-.02	-.49	58	67M
2. Y-Y-N	3	-2.71	-3.03	-.48	-.1.06	38	45M
	4	-2.69	-2.50	.37	-.08	38	56M
	5	-2.44	-2.52	-.24	-.66	38	67M
	4	-3.29	-3.08	.68	-.14	52	45M
	5	-3.04	-2.42	1.71	1.30	52	56M
	6	-3.06	-2.85	1.11	.52	52	67M
	5	-2.96	-2.45	1.36	.55	39	45M
	6	-2.86	-2.21	2.43	2.11	39	56M
	7	-1.78	-2.31	-3.26	-3.41	39	67M
	6	-2.68	-2.66	.38	-.28	32	45M
	7	-2.45	-2.49	-.25	-.71	32	56M
	8	-2.24	-1.97	2.20	1.69	32	67M
3. Y-N-Y	3	-1.82	-1.99	-.26	-.68	14	45M
	4	-1.33	-2.63	-3.05	-3.53	14	56M
	5	-2.45	-2.23	.80	.33	14	67M
	4	-2.25	-2.04	.67	.07	29	45M
	5	-2.25	-3.13	-2.31	-2.91	29	56M
	6	-2.46	-2.77	-1.52	-1.87	29	67M
	5	-2.29	-2.11	.49	-.10	27	45M
	6	-1.86	-2.29	-1.60	-2.06	27	56M
	7	-2.25	-2.28	-.06	-.46	27	67M
	6	-3.50	-3.10	2.10	1.10	33	45M
	7	-3.30	-3.71	-1.49	-2.22	33	56M
	8	-3.05	-2.97	.78	.21	33	67M
4. N-Y-Y	3	-2.97	-3.28	-.46	-1.10	47	45M
	4	-3.38	-2.85	1.17	.65	47	56M
	5	-3.17	-2.57	2.16	1.50	47	67M
	4	-2.77	-2.96	-.45	-1.04	37	45M
	5	-3.88	-3.14	2.03	1.48	37	56M
	6	-3.32	-2.83	2.55	1.84	37	67M
	5	-3.01	-2.80	.59	-.16	44	45M
	6	-2.81	-2.74	.28	-.21	44	56M
	7	-2.68	-2.74	-.28	-.73	44	67M
	6	-2.80	-2.67	.85	.12	25	45M
	7	-2.91	-2.99	-.41	-.97	25	56M
	8	-2.74	-2.65	.84	.32	25	67M

REFERENCES

Adams, M., Ingebo, G., Leitner, D., & Miller, P. (1983) Evaluation report on HOSTS program in the Portland Public Schools. Portland, OR: Portland Public Schools, Research and Evaluation Department. (ERIC Document Reproduction Services No. ED 241 578)

Adams, M., Ingebo, G., Leitner, D., Miller, S., Moilanen, C., & Peters, L. (1983). Computer use in the Portland Public Schools, an evaluation: computer-based instruction, computer literacy programs, computer management. Portland, OR: Portland Public Schools, Research and Evaluation Department. (ERIC Document Reproduction Services No. ED 240 158)

Claus, Richard N., & Girrbach, Charmaine, J. (1982). Long-term sustained effects of Chapter I participation, 1979-1982. Saginaw, MI: Saginaw Public Schools, Michigan Department of Evaluation Services. (ERIC Document Reproduction Service No. ED 227 147)

DeVito, Pasquale J., & Rubinstein, Sherry S. (1977, May) A follow-up study of Rhode Island Title I participants. Paper presented at the annual meeting of the New England Educational Research Organization, Manchester, NH (ERIC Document Reproduction Service No. ED 173 349)

Good, T. L. & Beckerman, T. M. (1978). Time on task: A naturalistic study in sixth grade classrooms. The Elementary School Journal, 78, 193-201.

Holmes, J. N. (1985). HOSTS pilot project: an evaluation report of HOSTS tutoring procedures for students at Cleveland High School during the 1984-85 school year. Portland, OR: Portland Public Schools, Research and Evaluation Department. (ERIC Document Reproduction Services No. ED 263 160)

Ingebo, G., Yagi, K., Leitner, D., Miller, S., & Miller, P. (1982). Evaluation report on three new instruction programs: Help One Student to Succeed, Prescription Learning, computer-assisted instruction. Portland, OR: Portland Public Schools, Research and Evaluation Department. (ERIC Document Reproduction Services No. ED 234 088)

Kiesling, H. (1975). The relationship of time spent on reading instruction to reading gains as measured by norm-referenced and criterion-referenced tests. Bloomington, IN: Indiana University. (ERIC Document Reproduction Services No. ED 116 135)

Kushmuk, J. (1985a). HOSTS program in the Portland Public Schools. 1984-85 evaluation report. Portland, OR: Portland Public Schools, Research and Evaluation Department. (ERIC Document Reproduction Services No. ED 268 167)

Kushmuk, J. (1985b). Prescription Learning in the Portland Public Schools. 1984-85 evaluation report. (Available from Portland Public Schools, Research and Evaluation Department, P.O. Box 3107, Portland, OR, 97208)

Leitner, D. & Ingebo, G. (1984a) HOSTS program in the Portland Public Schools. 1983-84 evaluation report. (Available from Portland Public Schools, Research and Evaluation Department, P.O. Box 3107, Portland, OR, 97208)

Leitner, D. & Ingebo, G. (1984b) Prescription Learning in the Portland Public Schools. 1983-84 evaluation report. Portland, OR: Portland Public Schools, Research and Evaluation Department. (ERIC Document Reproduction Services No. ED 251 490)

Whitmer, Jean E. (1986, Oct). Does writing modeled after children's picture books improve reading comprehension? Paper presented at the annual reading conference of the University of Colorado, Greeley, CO. (ERIC Document Reproduction Services No. ED 273 943)

Wyne, M. D. & Stuck, G. B. (1979). Time-on-task and reading performance in underachieving children. Journal of Reading Behavior, 11, 119-128.

Yagi, K. (1984, April). Chapter 1 evaluation - a strategy and results. Paper presented at the meeting of the Washington Educational Research Association, Seattle WA.

Yap, Kim Onn. (1987, April) Improving Chapter 1 through parents: a family goal program. Paper presented at the annual meeting of the American Educational Research Association, Washington, D. C. (ERIC Document Reproduction Services No. ED 285 895)

APPENDIX A

Comparisons of Predicted with Obtained

Results for Reading Participation Sequences--Following are the results showing the extent of agreement between those predicted and obtained for each reading group:

READING GROUP	AGREEMENT WITH PREDICTION
Y-Y-Y	Predicted lowest Fall mean; gain equal to District average and increasing slightly over years two and three. 1983-86: **High** 1984-87: **High** Obtained: Lowest Fall mean as predicted. Slight upward trend in 1983-86 compared to slight downward trend in 1984-87. Grade sequence 6-7-8 had lower than predicted gain in second year.
Y-Y-N	Predicted second lowest Fall mean; District average and increase second year, slight drop in third year. 1983-86: **Moderate** 1984-87: **Low** Obtained: Almost tied for second lowest Fall mean. Slight upward trend then drop in 1983-86, slight downward trend then drop in 1984-87. Larger drop than expected in third years.
Y-N-Y	Predicted third lowest Fall mean; U-shaped graph with above average gains in first and third years, below the second year.

1983-86: **High**
1984-87: **High**

Obtained: Almost tied for second lowest Fall mean. Nearly as predicted with pattern more prominent in 1984-87. Grade sequence 4-5-6 showed lower than expected gain in third year.

N-Y-Y Predicted fourth lowest Fall mean; below District gain the first year, near average the second and above average gain the third.

1983-86: **Moderate**
1984-87: **High**

Obtained: Fourth lowest Fall mean as predicted. Larger than expected gain the second year and tailing off the third year for 1983-86 group, 1984-87 group pattern nearly as predicted.

Y-N-N Predicted higher Fall mean than two or three year Chapter 1 students, but lowest of one year students; above District average gain the first year, declines for second and third years.

1983-86: **High**
1984-87: **Moderate**

Obtained: Fifth lowest Fall mean as predicted. Pattern as predicted for 1983-86 group, 1984-87 group pattern more level for second and third years. Grade sequence 6-7-8 did not fit pattern in both groups.

N-Y-N Predicted sixth lowest, second highest, Fall mean; lower than average gain the first year, large gain the second, and average the third.

1983-1986: **Moderate**
1984-1987: **Moderate**

Obtained: Sixth lowest, second highest Fall mean as predicted but essentially the same as the N-N-Y group. Patterns for both three-year groups much as predicted except smaller than expected gains the third year especially for 1983-86 group.

N-N-Y Predicted highest Fall mean; average or below gain the first year, lower the second, and above average gain the third.

1983-86: **High**
1984-87: **High**

Obtained: Highest fall mean as predicted but essentially the same as the N-Y-N group. Patterns for both three year groups much as predicted.

Results for Math Participation Sequences--Math achievement gain patterns were expected to be much the same as those for reading. Following are the results showing the extent of agreement between those predicted and obtained for each math group:

MATH GROUP AGREEMENT WITH PREDICTION

Y-Y-Y Predicted lowest Fall mean; District average gain increasing slightly years two and three.

1983-1986: **Moderate**
1984-1987: **Moderate**

Obtained: Lowest Fall mean as predicted. Gain as predicted but slightly larger than expected the first two years. Return to a smaller gain the third year was not predicted.

Y-Y-N Predicted second lowest Fall mean; District average and increasing second year, slight drop in third year.

1983-86: **Low**
1984-87: **Moderate**

Obtained: Essentially tied for second lowest Fall mean. District average to above average gains the first two years with unexpectedly large drop the third year for the 1983-86 group and mixed for the 1984-87 group.

Y-N-Y Predicted third lowest Fall mean; U-shaped graph with above average gains the first and third years, below the second.

1983-84: **Moderate**
1984-87: **Moderate**

Obtained: Fall mean about fifth lowest or higher than expected. Graph is somewhat U-shaped but not as high as predicted for the third year. Grade sequence 5-6-7 for 1983-86 showed a flatter pattern than predicted.

N-Y-Y Predicted fourth lowest Fall mean; below District gain the first year, near average the second, and above average gain the third.

1983-86: **High**
1984-87: **Moderate**

Obtained: Essentially tied for second or third lowest Fall mean. Gains were much as predicted for 1983-86, but gain patterns for grade sequences 5-6-7 and 6-7-8 for 1984-87 were flatter than predicted.

Y-N-N Predicted higher Fall mean than tow or three year Chapter 1 students, but lowest of one year students ; above District average gain the first year, declines for second and third years.

1983-86: **High**
1984-87: **Moderate**

Obtained: Fall mean was lowest among the one year groups as predicted, but about equal or lower than one of the two year groups, Y-N-Y. Gain pattern about as predicted for 1983-86, but gains for grade sequences 3-4-5 and 6-7-8 for 1984-87 were larger than expected in the third year.

N-Y-N Predicted second highest Fall mean; lower than average gain the first year, large gain the second, and average the third year.

1983-86: **Low**
1984-87: **Low**

Obtained: Fall mean was the highest by a slight margin. Gain patterns were as predicted the first two years for both groups, but the gains for the third year were smaller than expected except for the 4-5-6 grade sequence for 1984-87.

N-N-Y Predicted highest Fall mean; average or below gain the first year, lower the second, and above average gain the third.

1983-86: **High**
1984-87: **High**

Obtained: Fall mean was only slightly lower than that for the N-Y-N group. Gain patterns were nearly as predicted for both three year groups. Grade sequence 5-6-7 for 1984-87 had a smaller gain than expected the third year.

Published by
Publication Services
Portland Public Schools
Portland, Oregon